



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/00, A01K 67/027, C07K 16/00</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 98/24893</b> <b>(43) International Publication Date:</b> 11 June 1998 (11.06.98)
<b>(21) International Application Number:</b> PCT/US97/23091 <b>(22) International Filing Date:</b> 3 December 1997 (03.12.97)  <b>(30) Priority Data:</b> 08/759,620 3 December 1996 (03.12.96) US  <b>(71) Applicant:</b> ABGENIX, INC. [US/US]; 7601 Dumbarton Circle, Fremont, CA 94555 (US).  <b>(72) Inventors:</b> JAKOBOVITS, Aya; 2021 Monterey Avenue, Menlo Park, CA 94025 (US). KUCHERLAPATI, Raju; 8 Gracie Lane, Darien, CT 06820 (US). KLAPHOLZ, Sue; 76 Peter Coutts Circle, Stanford, CA 94305 (US). MENDEZ, Michael; P.O. Box 787, El Granada, CA 94018 (US). GREEN, Larry; 70 Crestline Drive #2, San Francisco, CA 94131 (US).  <b>(74) Agents:</b> HALEY, James, F., Jr. et al.; Fish & Neave, 1251 Avenue of the Americas, New York, NY 10020 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> TRANSGENIC MAMMALS HAVING HUMAN IG LOCI INCLUDING PLURAL V <sub>H</sub> AND V <sub>κ</sub> REGIONS AND ANTIBODIES PRODUCED THEREFROM  <b>(57) Abstract</b> <p>The present invention relates to transgenic non-human animals that are engineered to contain human immunoglobulin gene loci. In particular, animals in accordance with the invention possess human Ig loci that include plural variable (V<sub>H</sub> and V<sub>κ</sub>) gene regions. Advantageously, the inclusion of plural variable region genes enhances the specificity and diversity of human antibodies produced by the animal. Further, the inclusion of such regions enhances and reconstitutes B-cell development to the animals, such that the animals possess abundant mature B-cells secreting extremely high affinity antibodies.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**TRANSGENIC MAMMALS HAVING HUMAN IG LOCI INCLUDING PLURAL**  
**V<sub>H</sub> AND V<sub>K</sub> REGIONS AND ANTIBODIES PRODUCED THEREFROM**

**FIELD OF THE INVENTION**

The present invention relates to transgenic non-human animals that are engineered to contain human immunoglobulin gene loci. In particular, animals in accordance with the invention possess human Ig loci that include plural variable (V<sub>H</sub> and V<sub>K</sub>) gene regions. Advantageously, the inclusion of plural variable region genes enhances the specificity and diversity of human antibodies produced by the animal. Further, the inclusion of such regions enhances and reconstitutes B-cell development to the animals, such that the animals possess abundant mature B-cells secreting extremely high affinity antibodies.

**BACKGROUND OF THE TECHNOLOGY**

The ability to clone and reconstruct megabase-sized human loci in YACs and to introduce them into the mouse germline provides a powerful approach to elucidating the functional components of very large or crudely mapped loci as well as generating useful models of human disease. Furthermore, the utilization of such technology for substitution of mouse loci with their human equivalents could provide unique insights into the expression and regulation of human gene products during development, their communication with other systems, and their involvement in disease induction and progression.

An important practical application of such a strategy is the "humanization" of the mouse humoral immune system. Introduction of human immunoglobulin (Ig) loci into mice in which the endogenous Ig genes have been inactivated offers the opportunity to study of the mechanisms underlying programmed expression and assembly of antibodies as well as their role in B-cell development. Furthermore, such a strategy could provide an ideal source for production of fully human monoclonal antibodies (Mabs) - an important milestone towards fulfilling the promise of antibody therapy in human disease. Fully human antibodies are expected to minimize the immunogenic and allergic responses intrinsic to mouse or mouse-derivatized Mabs and thus to increase the efficacy and safety of the administered antibodies. The use of fully human antibodies can be expected to provide a substantial advantage in the treatment of chronic and recurring human diseases, such as inflammation, autoimmunity, and cancer, which require repeated antibody administrations.

One approach towards this goal was to engineer mouse strains deficient in mouse antibody production with large fragments of the human Ig loci in anticipation that such mice would produce a large repertoire of human antibodies in the absence of mouse antibodies. Large human Ig fragments would preserve the large variable gene diversity as well as the proper regulation of antibody production and expression. By exploiting the mouse machinery for antibody diversification and selection and the lack of immunological tolerance to human proteins, the reproduced human antibody repertoire in these mouse strains should yield high affinity antibodies against any antigen of interest, including human antigens. Using the hybridoma technology, antigen-specific human Mabs with the desired specificity could be readily produced and selected.

This general strategy was demonstrated in connection with our generation of the first XenoMouse™ strains as published in 1994. See Green et al. *Nature Genetics* 7:13-21 (1994). The XenoMouse™ strains were engineered with 245 kb and 190 kb-sized germline configuration fragments of the human heavy chain loci and kappa light chain loci, respectively, which contained core variable and constant region sequences. *Id.* The human Ig containing yeast artificial chromosomes (YACs) proved to be compatible with the mouse system for both rearrangement and expression of antibodies, and were capable of substituting for the inactivated mouse Ig genes. This was demonstrated by their ability to induce B-cell development and to produce an adult-like human



repertoire of fully human antibodies and to generate antigen-specific human Mabs. These results also suggested that introduction of larger portions of the human Ig loci containing greater numbers of V genes, additional regulatory elements, and human Ig constant regions might recapitulate substantially the full repertoire that is characteristic of the human humoral response to infection and immunization.

5        Such approach is further discussed and delineated in U.S. Patent Application Serial Nos. 07/466,008, filed January 12, 1990, 07/610,515, filed November 8, 1990, 07/919,297, filed July 24, 1992, 07/922,649, filed July 30, 1992, filed 08/031,801, filed March 15, 1993, 08/112,848, filed August 27, 1993, 08/234,145, filed April 28, 1994, 08/376,279, filed January 20, 1995, 08/430, 938, April 27, 1995, 08/464,584, filed June 5, 1995, 08/464,582, filed June 5, 1995, 08/463,191, filed June  
10    5, 1995, 08/462,837, filed June 5, 1995, 08/486,853, filed June 5, 1995, 08/486,857, filed June 5, 1995, 08/486,859, filed June 5, 1995, 08/462,513, filed June 5, 1995, and 08/724,752, filed October 2, 1996. *See also* European Patent No., EP 0 463 151 B1, grant published June 12, 1996, International Patent Application No., WO 94/02602, published February 3, 1994, International Patent Application No., WO 96/34096, published October 31, 1996, and PCT Application No.  
15    PCT/US96/05928, filed April 29, 1996. The disclosures of each of the above-cited patents and applications are hereby incorporated by reference in their entirety.

In an alternative approach, others, including GenPharm International, Inc., have utilized a "minilocus" approach. In the minilocus approach, an exogenous Ig locus is mimicked through the inclusion of pieces (individual genes) from the Ig locus. Thus, one or more V<sub>H</sub> genes, one or more  
20    D<sub>H</sub> genes, one or more J<sub>H</sub> genes, a mu constant region, and a second constant region (preferably a gamma constant region) are formed into a construct for insertion into an animal. This approach is described in U.S. Patent No. 5,545,807 to Surani et al. and U.S. Patent Nos. 5,545,806 and 5,625,825, both to Lonberg and Kay, and GenPharm International U.S. Patent Application Serial Nos. 07/574,748, filed August 29, 1990, 07/575,962, filed August 31, 1990, 07/810,279, filed  
25    December 17, 1991, 07/853,408, filed March 18, 1992, 07/904,068, filed June 23, 1992, 07/990,860, filed December 16, 1992, 08/053,131, filed April 26, 1993, 08/096,762, filed July 22, 1993, 08/155,301, filed November 18, 1993, 08/161,739, filed December 3, 1993, 08/165,699, filed December 10, 1993, 08/209,741, filed March 9, 1994, the disclosures of which are hereby incorporated by reference. *See also* International Patent Application Nos. WO 94/25585, published

November 10, 1994, WO 93/12227, published June 24, 1993, WO 92/22645, published December 23, 1992, WO 92/03918, published March 19, 1992, the disclosures of which are hereby incorporated by reference in their entirety. *See further* Taylor et al., 1992, Chen et al., 1993, Tuaillon et al., 1993, Choi et al., 1993, Lonberg et al., (1994), Taylor et al., (1994), and Tuaillon et al., (1995), the disclosures of which are hereby incorporated by reference in their entirety.

The inventors of Surani et al., cited above, and assigned to the Medical Research Counsel (the "MRC"), produced a transgenic mouse possessing an Ig locus through use of the minilocus approach. The inventors on the GenPharm International work, cited above, Lonberg and Kay, following the lead of the present inventors, proposed inactivation of the endogenous mouse Ig locus coupled with substantial duplication of the Surani et al. work.

An advantage of the minilocus approach is the rapidity with which constructs including portions of the Ig locus can be generated and introduced into animals. Commensurately, however, a significant disadvantage of the minilocus approach is that, in theory, insufficient diversity is introduced through the inclusion of small numbers of V, D, and J genes. Indeed, the published work appears to support this concern. B-cell development and antibody production of animals produced through use of the minilocus approach appear stunted. Therefore, the present inventors have consistently urged introduction of large portions of the Ig locus in order to achieve greater diversity and in an effort to reconstitute the immune repertoire of the animals.

Accordingly, it would be desirable to provide transgenic animals containing more complete germline sequences and configuration of the human Ig locus. It would be additionally desirable to provide such locus against a knockout background of endogenous Ig.

#### SUMMARY OF THE INVENTION

Provided in accordance with the present invention are transgenic animals having a near complete human Ig locus, including both a human heavy chain locus and a human kappa light chain locus. Preferably, the heavy chain locus includes greater than about 20%, more preferably greater than about 40%, more preferably greater than about 50%, and even more preferably greater than about 60% of the human heavy chain variable region. In connection with the human kappa light chain, preferably, the locus includes greater than about 20%, more preferably greater than about 40%,

more preferably greater than about 50%, and even more preferably greater than about 60% of the human kappa light chain variable region. Such percentages preferably refer to percentages of functional variable region genes.

Further, preferably such animals include the entire D<sub>H</sub> region, the entire J<sub>H</sub> region, the human mu constant region, and can additionally be equipped with genes encoding other human constant regions for the generation of additional isotypes. Such isotypes can include genes encoding  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\alpha$ ,  $\epsilon$ ,  $\beta$ , and other constant region encoding genes. Alternative constant regions can be included on the same transgene, i.e., downstream from the human mu constant region, or, alternatively, such other constant regions can be included on another chromosome. It will be appreciated that where such other constant regions are included on the same chromosome as the chromosome including the human mu constant region encoding transgene, cis-switching to the other isotype or isotypes can be accomplished. On the other hand, where such other constant region is included on a different chromosome from the chromosome containing the mu constant region encoding transgene, trans-switching to the other isotype or isotypes can be accomplished. Such arrangement allows tremendous flexibility in the design and construction of mice for the generation of antibodies to a wide array of antigens.

Preferably, such mice additionally do not produce functional endogenous immunoglobulins. This is accomplished in a preferred embodiment through the inactivation (or knocking out) of endogenous heavy and light chain loci. For example, in a preferred embodiment, the mouse heavy chain J-region and mouse kappa light chain J-region and C<sub>κ</sub>-region are inactivated through utilization of homologous recombination vectors that replace or delete the region. Such techniques are described in detail in our earlier applications and publications.

Unexpectedly, transgenic mice in accordance with the invention appear to possess an almost entirely reconstituted immune system repertoire. This is dramatically demonstrated when four separate mouse strains are compared: a first strain contains extensive human heavy chain variable regions and human kappa light chain variable regions and encodes only a mu isotype, a second strain contains extensive human heavy chain variable regions and human kappa light chain variable regions and encodes a mu and gamma-2 isotypes, a third strain contains significantly less human heavy and kappa light chain variable regions, and a fourth strain contains a double-inactivated mouse Ig locus.

The first and second strains undergo similar, if not identical, B-cell development, whereas the third strain has a reduced development and maturation of B-cells, and the fourth strain contains no mature B-cells. Further, it is interesting to note that production of human antibodies in preference to mouse antibodies is substantially elevated in mice having a knock-out background of endogenous Ig. That is to say that mice that contain a human Ig locus and a functionally inactivated endogenous Ig produce human antibodies at a rate of approximately 100 to 1000 fold as efficiently as mice that contain only a human Ig locus.

Thus, in accordance with a first aspect of the present invention there is provided a transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising: an inactivated endogenous immunoglobulin (Ig) locus, such that the mammal would not display normal B-cell development; an inserted human heavy chain Ig locus in substantially germline configuration, the human heavy chain Ig locus comprising a human mu constant region and regulatory and switch sequences thereto, a plurality of human  $J_H$  genes, a plurality of human  $D_H$  genes, and a plurality of human  $V_H$  genes; and an inserted human kappa light chain Ig locus in substantially germline configuration, the human kappa light chain Ig locus comprising a human kappa constant region, a plurality of  $J_k$  genes, and a plurality of  $V_k$  genes, wherein the number of  $V_H$  and  $V_k$  genes inserted are selected to substantially restore normal B-cell development in the mammal. In a preferred embodiment, the heavy chain Ig locus comprises a second constant region selected from the group consisting of human gamma-1, human gamma-2, human gamma-3, human gamma-4, alpha, delta, and epsilon. In another preferred embodiment, the number of  $V_H$  genes is greater than about 20. In another preferred embodiment, the number of  $V_k$  genes is greater than about 15. In another preferred embodiment, the number of  $D_H$  genes is greater than about 25, the number of  $J_H$  genes is greater than about 4, the number of  $V_H$  genes is greater than about 20, the number of  $J_k$  genes is greater than about 4, and the number of  $V_k$  genes is greater than about 15. In another preferred embodiment, the number of  $D_H$  genes, the number of  $J_H$  genes, the number of  $V_H$  genes, the number of  $J_k$  genes, and the number of  $V_k$  genes are selected such that the Ig loci are capable of encoding greater than about  $1 \times 10^5$  different functional antibody sequence combinations. In a preferred embodiment, in a population of mammals B-cell function is reconstituted on average to greater than about 50% as compared to wild type.

In accordance with a second aspect of the present invention there is provided an improved transgenic non-human mammal having a genome that comprises modifications, the modifications rendering the mammal capable of producing human immunoglobulin molecules but substantially incapable of producing functional endogenous immunoglobulin molecules, the improvement comprising: insertion into the genome of the mammal of sufficient human  $V_H$ ,  $D_H$ ,  $J_H$ ,  $V\kappa$ , and  $J\kappa$  genes such that the mammal is capable encoding greater than about  $1 \times 10^6$  different functional human immunoglobulin sequence combinations.

In accordance with a third aspect of the present invention, there is provided an improved transgenic non-human mammal having a genome that comprises modifications, the modifications rendering the mammal capable of producing human immunoglobulin molecules but substantially incapable of producing functional endogenous immunoglobulin molecules, which modifications, with respect to the mammal's incapacity to produce functional endogenous immunoglobulin molecules would not allow the mammal to display normal B-cell development, the improvement comprising: insertion into the genome of the mammal of sufficient human  $V_H$ ,  $D_H$ ,  $J_H$ ,  $V\kappa$ , and  $J\kappa$  genes such that the mammal is capable of encoding greater than about  $1 \times 10^6$  different functional human immunoglobulin sequence combinations and sufficient  $V_H$  and  $V\kappa$  genes to substantially restore normal B-cell development in the mammal. In a preferred embodiment, in a population of mammals B-cell function is reconstituted on average to greater than about 50% as compared to wild type.

In accordance with a fourth aspect of the present invention, there is provided a transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising: an inactivated endogenous heavy chain immunoglobulin (Ig) locus; an inactivated endogenous kappa light chain Ig locus; an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of  $\gamma H2$ ; and an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of  $\gamma K2$ .

In accordance with a fifth aspect of the present invention there is provided a transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising: an inactivated endogenous heavy chain immunoglobulin (Ig) locus; an inserted human

heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2; and an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

5 In accordance with a sixth aspect of the present invention, there is provided a transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising: an inactivated endogenous heavy chain immunoglobulin (Ig) locus; an inactivated endogenous kappa light chain Ig locus; an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide  
10 sequence of yH2 without the presence of a human gamma-2 constant region; and an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

In accordance with a seventh aspect of the present invention, there is provided a  
transgenic non-human mammal having a genome, the genome comprising modifications, the  
15 modifications comprising: an inactivated endogenous heavy chain immunoglobulin (Ig) locus; an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2 without the presence of a human gamma-2 constant region; and an inserted human kappa light chain Ig locus, the human kappa  
light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide  
20 sequence of yK2.

In accordance with an eighth aspect of the present invention, there is provided a method for the production of human antibodies, comprising: inoculating any of the mammals of the first through fifth aspects of the invention discussed above with an antigen; collecting and immortalizing lymphocytic cells to obtain an immortal cell population secreting human antibodies that specifically  
25 bind to the antigen with an affinity of greater than  $10^9 \text{ M}^{-1}$ ; and isolating the antibodies from the immortal cell populations.

In a preferred embodiment, the antigen is IL-8. In another preferred embodiment, the antigen is EGFR. In another preferred embodiment, the antigen is TNF- $\alpha$ .

In accordance with a ninth aspect of the present invention, there is provided an antibody

produced by the method of the sixth aspect of the invention, including antibodies to IL-8, EGFR, and TNF- $\alpha$ .

In accordance with a tenth aspect of the present invention, there is provided an improved method for the production of transgenic mice, the transgenic mice having a genome, the genome comprising modifications, the modifications comprising insertion of a plurality of human variable regions, the improvement comprising: insertion of the human variable regions from a yeast artificial chromosome.

In accordance with an eleventh aspect of the present invention, there are provided transgenic mice and transgenic offspring therefrom produced through use of the improvement of the eighth aspect of the present invention.

In accordance with a twelfth aspect of the present invention, there is provided a transgenic mammal, the transgenic mammal comprising a genome, the genome comprising modifications, the modifications comprising an inserted human heavy chain immunoglobulin transgene, the improvement comprising: the transgene comprising selected sets of human variable region genes that enable human-like junctional diversity and human-like complementarity determining region 3 (CDR3) lengths. In a preferred embodiment, the human-like junctional diversity comprises average N-addition lengths of 7.7 bases. In another preferred embodiment, the human-like CDR3 lengths comprise between about 2 through about 25 residues with an average of about 14 residues.

#### BRIEF DESCRIPTION OF THE DRAWING FIGURES

**Figure 1** is a schematic representation of the reconstructed human heavy chain and human kappa light chain loci YACs introduced into preferred mice in accordance with the invention. YACs spanning the human heavy chain (1H, 2H, 3H, and 4H) and the human kappa light chain proximal (1K, 2K, and 3K) loci were cloned from human -YAC libraries. The locations of the different YACs with respect to the human Ig loci (adopted from Cook and Tomlinson, 1995, and Cox et al., 1994), their sizes, and non-Ig sequences are indicated (not shown to scale). The YACs were recombined into yeast in a two-step procedure (*see* Materials and Methods) to reconstruct the human heavy and kappa light chain YACs. yH2, the human heavy chain containing YAC, was further retrofitted with a human  $\gamma_2$  gene sequence. yK2, was the human kappa light chain containing YAC. The YAC vector

elements: telomere ▲, centromere ●, mammalian (HPRT, Neo) and yeast selectable markers (TRP1, ADE2, LYS2, LEU2, URA3, HIS3) on the YAC vector arms are indicated.  $V_H$  segments are classified as genes with open reading frame ●, pseudogenes □, and unsequenced genes ○.  $V_k$  segments are classified as genes with open reading frames ●, and pseudogenes □. The V genes that we have found to be utilized by the XenoMouse II are marked (\*). The  $V_H$  gene region contained on yH2 is marked by arrows.

**Figure 2** is a series of Southern Blot analyses and characterizations of the human heavy chain YAC, yH2, integrated in ES cells and in XenoMouse strains. Figure 2a is a series of Southern Blot analyses of EcoRI (a, c) and BamHI (b, d, e) digested DNA (2 µg) prepared from the CGM1 immortalized B-lymphoblast cell line derived from the Washington University YAC library source (Brownstein et al., 1989), yH2 YAC (0.5 µg YAC added to 2 µg of 3B1 DNA), unmodified E14TG.3B1 (3B1), and yH2-containing ES cell lines: L10, J9.2, L18, L17, and J17. The probes used for blotting were human  $V_H1$  (a),  $D_H$  (b) [18 kb fragment in CGM1 lane represents D segments on chromosome 16],  $V_H3$  (c),  $C_\mu$  (d) and  $J_H$  (e). Figure 2b is a series of Southern Blot analyses of EcoRI (a-b) and BamHI (c-d) digested DNA (10 µg) that was prepared from the tails of wildtype (WT, 129xB57BL/6J), XM2A-1, and XM2A-2 (2 individual offspring) mice or from the parental yH2-containing ES cell lines L10 (slightly underloaded relative to other samples), J9.2, and yK2-containing ES cell line J23.1. The probes used were human  $V_H1$  (a),  $V_H4$  (b), human  $\gamma$ -2 (c), and mouse 3'-enhancer (d, the 5kb band represents the endogenous mouse 3'-enhancer fragment). Fragment sizes of molecular weight markers (in kb) are indicated.

**Figure 3** is a series of Southern Blot analyses characterizing the human kappa light chain YAC, yK2, integrated in ES cells and in XenoMouse 2A Strains. Figure 2a is a series of Southern Blot analyses of EcoRI (a, c, d) and BamHI (b, e) digested DNA (2 µg) prepared from CGM1 cell line (Brownstein et al., 1989, *supra*), yK2 YAC (0.5 µg YAC DNA added to 2 µg of 3B1 DNA), unmodified E14TG.3B1 (3B1), and yK2-containing ES cell lines: J23.1 and J23.7. The probes used were human  $V_k$  (a),  $K_d$  (b),  $V_kII$  (c),  $V_kIII$  (d), and  $C_k$  (e). Figure 2b is a series of Southern Blot analyses of EcoRI-digested DNA (2 µg) that was prepared from the tails of wildtype (WT, 129xB6), XM2A-1, and XM2A-2 (2 individual offspring) mice or from the parental yH2-containing ES cell lines L10 (slightly underloaded relative to other samples), J9.2, and yK2-containing ES cell line J23.1.



The probes that were used were human V<sub>κ</sub>I (a), V<sub>κ</sub>IV (b), V<sub>κ</sub>VI (c) and 3'-enhancer (d). Fragment sizes of molecular weight markers (in kb) are indicated.

**Figure 4** shows B-cell reconstitution and surface expression of human  $\mu$ ,  $\delta$ , and  $\kappa$  chains on XenoMouse-derived B-cells and shows flow cytometry analysis of peripheral blood (Fig. 4a) and spleen (Fig. 4b) lymphocytes from wildtype mice (WT), double inactivated mice (DI), and XenoMouse strains 2A-1 and 2A-2 (XM2A-1, XM2A-2). Four-color flow cytometry analysis was carried out using antibodies to the B-cell-specific marker B220 in combination with anti-human  $\mu$ ,  $\delta$ ,  $\kappa$ , or mouse  $\mu$ ,  $\delta$ ,  $\kappa$ , or  $\lambda$ . The percentage of positively-stained cells is shown in each quadrant. Isolation and staining of cells were performed as described in Materials and Methods. Populations of human  $\kappa^+$  and mouse  $\lambda^+$  cells were determined after first gating for B220<sup>+</sup> $\mu^+$  populations in the indicated region. Populations of  $\mu^+$  and  $\delta^+$  cells were determined after first gating for B220<sup>+</sup> cells. The percentage of positive cells within a region or quadrant is indicated. The FACS profiles shown are representative of several experiments performed on each of the strains.

**Figure 5** shows that XenoMouse-derived human antibodies block the binding of their specific antigens to cells. Figure 5a shows the inhibition of labeled [<sup>125</sup>I] IL-8 binding to human neutrophils by the mouse anti-human IL-8 antibody (R&D Systems) (□) and the fully human Mabs D1.1 (◆), K2.2 (●), K4.2 (▲), and K4.3 (▼). The background binding of labeled [<sup>125</sup>I]IL-8 in the absence of antibody was 2657 cpm. Figure 5b shows the inhibition of labeled [<sup>125</sup>I]EGF to its receptors on A431 cells by mouse anti-human EGFR antibodies 225 and 528 (□, ▽, respectively; Calbiochem) and the fully human antibodies E1.1(●), E2.4 (▲), E2.5 (▼) and E2.11 (◆). The background binding of [<sup>125</sup>I]EGF in the absence of antibodies was 1060 cpm. Figure 5c shows inhibition of labeled [<sup>125</sup>I] TNF- $\alpha$  binding to its receptors on U937 cells by the mouse anti-human TNF- $\alpha$  antibody (R&D Systems) (□) and fully human Mabs T22.1 (◆), T22.4 (●), T22.8 (▲), and T22.9 (■). The background binding of [<sup>125</sup>I]TNF- $\alpha$  in the absence of antibody was 4010 cpm. Control human IgG<sub>2</sub> myeloma antibody (⊗).

**Figure 6** shows repertoire and somatic hypermutation in XenoMouse-derived fully human Mabs. Predicted amino acid sequences of four anti-IL-8 (Fig. 6a) and four anti-EGFR (Fig. 6b) human IgG<sub>2</sub> $\kappa$  Mabs, divided into CDR1, CDR2 and CDR3 and the constant regions, C<sub>γ</sub>2 and C<sub>κ</sub>. The D and J genes of each antibody are indicated. The amino acid substitutions from the germline

sequences are indicated in bold letters.

**Figure 7** is a schematic diagram of the human heavy chain genome and the human kappa light chain genome.

**Figure 8** is another schematic diagram showing the construction of the yH2 (human heavy chain) YAC.

**Figure 9** is another schematic diagram showing the construction of the yK2 (human kappa light chain) YAC.

**Figure 10** is another schematic diagram showing the construction of the yK2 (human kappa light chain) YAC.

**Figure 11** shows a series of Southern Blot analyses demonstrating integration intact of the yH2 (human heavy chain) YAC into ES cells and into the mouse genome. Detailed discussion is provided in connection with Figure 2.

**Figure 12** shows a series of Southern Blot analyses demonstrating integration intact of the yK2 (human kappa light chain) YAC into ES cells and into the mouse genome. Detailed discussion is provided in connection with Figure 3.

**Figure 13** shows B-cell reconstitution and surface expression of human  $\mu$ ,  $\delta$ , and  $\kappa$  chains and mouse  $\lambda$  chains on XenoMouse-derived B-cells and shows flow cytometry analysis of peripheral blood. Further details are provided in connection with Figure 4.

**Figure 14** shows production levels of human antibodies by XenoMouse II strains in comparison to murine antibody production by wild type mice.

**Figure 15** is a repertoire analysis of human heavy chain transcripts expressed in XenoMouse II strains.

**Figure 16** is a repertoire analysis of human kappa light chain transcripts expressed in XenoMouse II strains.

**Figure 17** is another depiction of the diverse utilization of human  $V_H$  and  $V_K$  genes that have been observed as utilized in XenoMouse II strains.

**Figure 18** shows the titers of human antibody production in XenoMouse II strains.

**Figure 19** is a depiction of gene utilization of anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 20** shows heavy chain amino acid sequences of anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 21** shows kappa light chain amino acid sequences of anti-IL-8 antibodies derived from XenoMouse II strains.

5        **Figure 22** shows blockage of IL-8 binding to human neutrophils by human anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 23** shows inhibition of CD11b expression on human neutrophils by human anti-IL-8 antibodies derived from XenoMouse II strains.

10       **Figure 24** shows inhibition of IL-8 induced calcium influx by human anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 25** shows inhibition of IL-8 RB/293 chemotaxis by human anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 26** is a schematic diagram of a rabbit model of human IL-8 induced skin inflammation.

15       **Figure 27** shows the inhibition of human IL-8 induced skin inflammation in the rabbit model of Figure 26 with human anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 28** shows inhibition of angiogenesis of endothelial cells on a rat corneal pocket model by human anti-IL-8 antibodies derived from XenoMouse II strains.

**Figure 29** is a depiction of gene utilization of human anti-EGFR antibodies derived from XenoMouse II strains.

20       **Figure 30** shows heavy chain amino acid sequences of human anti-EGFR antibodies derived from XenoMouse II strains.

**Figure 31** shows blockage EGF binding to A431 cells by human anti-EGFR antibodies derived from XenoMouse II strains.

25       **Figure 32** shows inhibition of EGF binding to SW948 cells by human anti-EGFR antibodies derived from XenoMouse II strains.

**Figure 33** shows that human anti-EGFR antibodies derived from XenoMouse II strains inhibit growth of SW948 cells *in vitro*.

**Figure 34** shows inhibition of TNF- $\alpha$  binding to U937 cells through use of human anti-TNF- $\alpha$  antibodies derived from XenoMouse II strains.

Figure 35 shows kappa light chain amino acid sequences of human anti-EGFR antibodies derived from XenoMouse II strains.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Herein we describe the generation and characterization of several strains of mice containing substantially germline configuration megabase-sized human Ig loci. The present invention thus provides the first demonstration of reconstruction of the large and complex human Ig loci on YACs and the successful introduction of megabase-sized YACs into mice to functionally replace the corresponding mouse loci.

#### Mouse Strains

The following mouse strains are described and/or utilized herein:

**Double Inactivated (DI) Strain:** The DI strain of mice are mice that do not produce functional endogenous, mouse, Ig. In preferred embodiments, the DI mice possess an inactivated mouse J<sub>H</sub> region and an inactivated mouse C<sub>κ</sub> region. The construction of this strain is discussed extensively elsewhere. For example, the techniques utilized for generation of the DI strains are described in detail in U.S. Patent Application Serial Nos. 07/466,008, filed January 12, 1990, 07/610,515, filed November 8, 1990, 07/919,297, filed July 24, 1992, 08/031,801, filed March 15, 1993, 08/112,848, filed August 27, 1993, 08/234,145, filed April 28, 1994, 08/724,752, filed October 2, 1996. See also European Patent No., EP 0 463 151 B1, grant published June 12, 1996, International Patent Application No., WO 94/02602, published February 3, 1994, International Patent Application No., WO 96/34096, published October 31, 1996, and PCT Application No. PCT/US96/05928, filed April 29, 1996. The disclosures of each of the above-cited patent and patent applications are hereby incorporated by reference in their entirety. It has been observed and reported that DI mice possess a very immature B-cell development. The mice do not produce mature B-cells, only pro-B-cells.

**XenoMouse I Strain:** The design, construction, and analysis of the XenoMouse I strain was discussed in detail in Green et al., *Nature Genetics*, 7:13-21 (1994). Such mice produced IgM $\kappa$  antibodies against a DI background. The mice showed improved B-cell function when compared to the DI strain of mice which have little to no B-cell development. While XenoMouse I strains of mice were capable of mounting a sizeable immune response to antigenic challenge, there appeared to be inefficient in their production of B-cells and possessed a limited response to different antigens which apparently was related to their limited V-gene repertoire.

**L6 Strain:** The L6 strain is a mouse producing IgM $\kappa$  antibodies against a DI background of endogenous mouse Ig. L6 mice contain an inserted human heavy chain and an inserted human kappa light chain. The L6 strain is generated through breeding of a mouse containing a heavy chain insert against a double inactivated background (L6H) and a mouse having a kappa light chain insert against a double inactivated background (L6L). The heavy chain insert comprises an intact approximately 970 kb human DNA insert from a YAC containing approximately 66 V<sub>H</sub> segments, starting at V<sub>H</sub>6-1 and ending at V<sub>H</sub>3-65, and including the major D gene clusters (approximately 32), J<sub>H</sub> genes (6), the intronic enhancer (Em), C $\mu$ , and through about 25 kb past C $\delta$ , in germline configuration. The light chain insert comprises an intact approximately 800 kb human DNA insert from a YAC which contains approximately 32 V $\kappa$  genes starting at V $\kappa$ -B3 and ending at V $\kappa$ -Op11. The 800 kb insert contains a deletion of approximately 100 kb starting at V $\kappa$ -Lp-13 and ending at V $\kappa$ -Lp-5. However, the DNA is in germline configuration from V $\kappa$ -Lp-13 to 100 kb past V $\kappa$ -Op-1, and also contains the J $\kappa$  genes, the intronic and 3' enhancers, the constant C $\kappa$  gene, and Kde. The L6H and L6L mice have been shown to access the full spectrum of the variable genes incorporated into their genome. It is expected that the L6 mice will similarly access the full spectrum of variable genes in their genome. Furthermore, L6 mice will exhibit predominant expression of human kappa light chain, a large population of mature B-cells, and normal levels of IgM $\kappa$  human antibodies. Such mice will mount a vigorous human antibody response to multiple immunogens, ultimately yielding antigen-specific fully human Mabs with subnanomolar affinities.

**XenoMouse IIa Strain:** The XenoMouse IIa mice represent our second generation

XenoMouse™ strains equipped with germline configuration megabase-sized human Ig loci, against a DI background, such that the mice do not produce functional endogenous Ig. Essentially, the mice are equivalent in construction to the L6 strain, but additionally include the human  $\gamma 2$  gene with its entire switch and regulatory sequences and the mouse 3' enhancer in cis. The mice contain an approximately 1020 kb heavy and an approximately 800 kb kappa light chain loci, reconstructed on YACs, which include the majority of the human variable region genes, including heavy chain genes (approximately 66  $V_H$ ) and kappa light chain genes (approximately 32  $V_K$ ), human heavy constant region genes ( $\mu$ ,  $\delta$ , and  $\gamma$ ) and kappa constant region genes ( $C_K$ ), and all of the major identified regulatory elements. These mice have been shown to access the full spectrum of the variable genes incorporated into their genome. Furthermore, they exhibit efficient class switching and somatic hypermutation, predominant expression of human kappa light chain, a large population of mature B-cells, and normal levels of IgM $_{\kappa}$  and IgG $_{\kappa}$  human antibodies. Such mice mount a vigorous human antibody response to multiple immunogens, including human IL-8, human EGF receptor (EGFR), and human tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), ultimately yielding antigen-specific fully human Mabs with subnanomolar affinities. This last result conclusively demonstrates XenoMouse™ as an excellent source for rapid isolation of high affinity, fully human therapeutic Mabs against a broad spectrum of antigens with any desired specificity.

As will be appreciated from the above-introduction, the XenoMouse II strain appears to undergo mature B-cell development and mount powerful adult-human-like immune responses to antigenic challenge. The L6 strain, as predicted from the data in connection with L6L and L6H mice, also appear to undergo mature B-cell development and mount powerful adult-human-like immune responses to antigenic challenge. When DI mice are compared to XenoMouse I strains and DI and XenoMouse I strains are compared to L6 and XenoMouse II strains, a markedly different B-cell development profile is observed. Owing to this difference, it appears that the quantity and/or quality of variable region sequences introduced into the animals are essential to the induction B-cell maturation and development and the generation of an adult-human-like immune response. Thus, in addition to the strains' clear use in the generation of human antibodies, the strains provide a valuable tool for studying the nature of human antibodies in the normal immune response, as well as the abnormal response characteristic of autoimmune disease and other disorders.

*Variable Region - Quantitative Diversity*

It is predicted that the specificity of antibodies (i.e., the ability to generate antibodies to a wide spectrum of antigens and indeed to a wide spectrum of independent epitopes thereon) is dependent upon the variable region genes on the heavy chain ( $V_H$ ) and kappa light chain ( $V_K$ ) genome. The human heavy chain genome includes approximately 95 functional genes which encode variable regions of the human heavy chain of immunoglobulin molecules. In addition, the human light chain genome includes approximately 40 genes on its proximal end which encode variable regions of the human kappa light chain of immunoglobulin molecules. We have demonstrated that the specificity of antibodies can be enhanced through the inclusion of a plurality of genes encoding variable light and heavy chains.

Provided in accordance with the present invention are transgenic mice having a substantial portion of the human Ig locus, preferably including both a human heavy chain locus and a human kappa light chain locus. In preferred embodiments, therefore, greater than 10% of the human  $V_H$  and  $V_K$  genes are utilized. More preferably, greater than about 20%, 30%, 40%, 50%, 60%, or even 70% or greater of  $V_H$  and  $V_K$  genes are utilized. In a preferred embodiment, constructs including 32 genes on the proximal region of the  $V_K$  light chain genome are utilized and 66 genes on the  $V_H$  portion of the genome are utilized. As will be appreciated, genes may be included either sequentially, i.e., in the order found in the human genome, or out of sequence, i.e., in an order other than that found in the human genome, or a combination thereof. Thus, by way of example, an entirely sequential portion of either the  $V_H$  or  $V_K$  genome can be utilized, or various V genes in either the  $V_H$  or  $V_K$  genome can be skipped while maintaining an overall sequential arrangement, or V genes within either the  $V_H$  or  $V_K$  genome can be reordered, and the like. In a preferred embodiment, the entire inserted locus is provided in substantially germline configuration as found in humans. In any case, it is expected and the results described herein demonstrate that the inclusion of a diverse array of genes from the  $V_H$  and  $V_K$  genome leads to enhanced antibody specificity and ultimately to enhanced antibody affinities.

Further, preferably such mice include the entire  $D_H$  region, the entire  $J_H$  region, the human mu constant region, and can additionally be equipped with other human constant regions for the coding and generation of additional isotypes of antibodies. Such isotypes can include genes encoding  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\gamma_4$ ,  $\alpha$ ,  $\epsilon$ , and  $\delta$  and other constant region encoding genes with appropriate switch and regulatory

sequences. As will be appreciated, and as discussed in more detail below, a variety of switch and regulatory sequences can be appropriately utilized in connection with any particular constant region selection.

The following Table indicates the diversity of antibody combinations that are possible in humans, based strictly on random V-D-J joining and combination with kappa light chains, without consideration of N-addition or somatic mutation events. Based on these considerations, there are greater than 3.8 million possible antibody combinations in humans, of any particular isotype.

**TABLE I**

Region	Heavy Chain	Kappa Light Chain
Variable "V"	~95	40
Diversity "D"	≥32	--
Joining "J"	6	5
Combinations (VxDxJ)	18,240	200
Total Combinations (HC Combinations x LC Combinations)	3.65 X 10 <sup>6</sup>	

In connection with a preferred embodiment of the invention, through the inclusion of about 66 V<sub>H</sub> genes and 32 V<sub>κ</sub> genes in a mouse with a full complement of D<sub>H</sub>, J<sub>H</sub>, and J<sub>κ</sub> genes, the possible diversity of antibody production is on the order of 2.03 X 10<sup>6</sup> different antibodies. As before, such calculation does not take into account N-addition or somatic mutation events. Therefore, it will be appreciated that mice in accordance with the invention, such as the L6 and the XenoMouse II strains, offer substantial antibody diversity. In preferred embodiments, mice are designed to have the capability of producing greater than 1 X 10<sup>6</sup> different heavy chain V-D-J combinations and kappa light chain V-J combinations, without accounting for N-additions or somatic mutation events.



Variable Region - Qualitative Diversity

In addition to quantitative diversity, quantitative selection of V-genes (i.e., large and diverse numbers of V-genes) and/or qualitative selection of V-genes (i.e., selection of particular V-genes) appears to play a role in what we refer to herein as “qualitative diversity.” Qualitative diversity, as used herein, refers to diversity in V-D-J rearrangements wherein junctional diversity and/or somatic mutation events are introduced. During heavy chain rearrangement, certain enzymes (RAG-1, RAG-2, and possibly others) are responsible for the cutting of the DNA representing the coding regions of the antibody genes. Terminal deoxynucleotidyl transferase (Tdt) activity is upregulated which is responsible for N-terminal additions of nucleotides between the V-D and D-Jgene segments. Similar enzymes and others (SCID and other DNA repair enzymes) are responsible for the deletion that occurs at the junctions of these coding segments. With respect to junctional diversity, both N-addition events and formation of the complementarity determining region 3 (CDR3), are included within such term. As will be appreciated, CDR3 is located across the D region and includes the V-D and D-J junctional events. Thus, N-additions and deletions during both D-J rearrangement and V-D rearrangement are responsible for CDR3 diversity.

It has been demonstrated that there are certain differences between murine and human junctional diversities. In particular, some researchers have reported that murine N-addition lengths and CDR3 lengths are generally shorter than typical human N-addition lengths and CDR3 lengths. Such groups have reported that, in humans, N-additions of about 7.7 bases in length, on average, are typically observed. Yamada et al. (1991). Mouse-like N-additions are more often on the order of about 3 bases in length, on average. Feeney et al. (1990). Similarly, human-like CDR3 lengths are longer than mouse-like CDR3's. In man CDR3 lengths of between 2 and 25 residues, with an average of 14 residues, is common. In mice, some groups have reported shorter average CDR3 lengths.

The junctional diversity created by N-additions and CDR3 additions play a clear role developing antibody specificity.

In accordance with the invention, rearranged V-D-J gene sequences show N-addition lengths that are comparable to expected adult-human N-addition lengths. Further, amino acid sequences across the open reading frame (ORF) corresponding to CDR3 sequences show CDR3 lengths that

are comparable to expected adult-human CDR3 lengths. Such data is indicative that quantitative variable region diversity and/or qualitative variable region diversity results in human-like junctional diversity. Such junctional diversity is expected to lead to a more human-like antibody specificity.

5 *Variable Region - Affinities*

While we have not conclusively demonstrated a direct causal connection between the increased variable region inclusion and antibody specificity, it appears, and it is expected that through providing such diversity, the ability of the mouse to mount an immune response to a wide array of antigens is possible and enhanced. Additionally, such mice appear more equipped to mount immune responses to a wide array of epitopes upon individual antigens or immunogens. From our data it also appears that antibodies produced in accordance with the present invention possess enhanced affinities. Such data includes comparisons between mice in accordance with the invention and the XenoMouse I strains, as well as consideration of the published results of GenPharm International and the MRC. In connection with the XenoMouse I strains, as mentioned above, such mice possessed inefficient B-cell production and a limited response to different antigens. Such result appeared related in part to the limited V-gene repertoire. Similarly, results reported by GenPharm International and the MRC indicate a limited response to diverse antigens.

Without wishing to bound to any particular theory or mode of operation of the invention, it would appear that enhanced affinities appear to result from the provision of the large number of V regions. From our data, the provision of greater numbers and/or selection of qualities of V-gene sequences, enhances junctional diversity (N-additions and formation of complementarity determining region 3 ("CDR3") diversity), which is typical of an adult-human-like immune response, and which play a substantial role in affinity maturation of antibodies. It may also be that such antibodies are more effective and efficient in somatic mutation events that lead to enhanced affinities. Each of junctional diversity and somatic mutation events are discussed in additional detail below.

With respect to affinities, antibody affinity rates and constants derived through utilization of plural  $V_H$  and  $V_K$  genes (i.e., the use of 32 genes on the proximal region of the  $V_K$  light chain genome and 66 genes on the  $V_H$  portion of the genome) results in association rates ( $k_a$  in  $M^{-1}S^{-1}$ ) of greater than about  $0.50 \times 10^{-6}$ , preferably greater than  $2.00 \times 10^{-6}$ , and more preferably greater than about

4.00 X 10<sup>-6</sup>; dissociation rates (kd in S<sup>-1</sup>) of greater than about 1.00 X 10<sup>-4</sup>, preferably greater than about 2.00 X 10<sup>-4</sup>, and more preferably greater than about 4.00 X 10<sup>-4</sup>; and dissociation constant (in M) of greater than about 1.00 X 10<sup>-10</sup>, preferably greater than about 2.00 X 10<sup>-10</sup>, and more preferably greater than about 4.00 X 10<sup>-10</sup>.

5 Preferably, such mice additionally do not produce functional endogenous immunoglobulins. This is accomplished in a preferred embodiment through the inactivation (or knocking out) of endogenous heavy and light chain loci. For example, in a preferred embodiment, the mouse heavy chain J-region and mouse kappa light chain J-region and C<sub>κ</sub>-region are inactivated through utilization of homologous recombination vectors that replace or delete the region.

#### Variable Region - B-cell Development

10 B-cell development is reviewed in Klaus *B Lymphocytes* (IRL Press (1990)) and Chapters 1-3 of *Immunoglobulin Genes* (Academic Press Ltd. (1989)), the disclosures of which are hereby incorporated by reference. Generally, in mammals, blood cell development, including B- and T-cell lymphocytes, originate from a common pluripotent stem cell. The lymphocytes, then, evolve from a common lymphoid progenitor cell. Following an early gestational period, B-cell initiation shifts from the liver to the bone marrow where it remains throughout the life of the mammal.

15 In the life cycle of a B-cell, the first generally recognizable cell is a pro-pre-B-cell which is found in the bone marrow. Such a cell has begun heavy chain V-D-J rearrangement, but does not yet make protein. The cell then evolves into a large, rapidly dividing, pre-B-cell I which is a cytoplasmically  $\mu^+$  cell. This pre-B-cell I then stops dividing, shrinks, and undergoes light chain V-J rearrangement becoming a pre-B-cell II which expresses surface IgM, which leave the marrow as immature B-cells. Most of the emerging immature B-cells continue to develop and to produce surface IgD, indicative of their completion of differentiation and development as fully mature immunocompetent peripheral B-cells, which reside primarily in the spleen. However, it is possible to eliminate the delta constant region and still obtain immunocompetent cells.

25 B-cell differentiation and development can be monitored and/or tracked through the use of surface markers. For example, the B220 antigen is expressed in relative abundance on mature B-cells in comparison to pre-B-cells I or II. Thus, cells that are B220<sup>+</sup> and surface IgM<sup>+</sup> ( $\mu^+$ ) can be utilized

to determine the presence of mature B-cells. Additionally, cells can be screened for surface IgD expression ( $\delta^+$ ). Another antigen, heat stable antigen, is expressed by pre-B-cells II as they transition to the periphery (i.e., as they become  $\mu^+$  and/or  $\mu^+$ ,  $\delta^+$ ).

5

**TABLE II**

	Bone Marrow			Spleen	
Marker	pro-pre-B-cell	pre-B-cell I	pre-B-cell II emerging B-cell	immature B-cell	mature B-cell
B220	-	-	$\pm$	+	++
HSA	-	-	+	$\pm$	-
$\mu$	-	-	+	+	+
$\delta^+$	-	-	-	-	+

10

15 \* Assuming the presence of a functional copy of the C $\delta$  gene on the transgene.

Through use of B-cell markers, such as those mentioned above, development and differentiation of B-cells can be monitored and assessed.

20 We have previously demonstrated that DI mice (mice that do not undergo heavy chain V-D-J rearrangement or light chain V-J rearrangement) do not produce mature B-cells. In fact, such mice arrest at the production of pro-pre-B-cells and B-cells never move from the bone marrow to peripheral tissues, including the spleen. Thus, both B-cell development and antibody production are completely arrested. The same result is seen in mice that are only heavy chain inactivated; B-cell  
25 development and differentiation arrests in the bone marrow.

Our XenoMouse I strain produced functional, somewhat mature B-cells. However, the numbers of B-cells, in both the bone marrow and peripheral tissues, were significantly reduced relative to wild type mice.

In contrast, our XenoMouse II strains and L6 strains, unexpectedly possess almost complete  
30 B-cell reconstitution. Therefore, in accordance with the invention, we have demonstrated that

through the quantitative inclusion or qualitative inclusion of variable region genes B-cell differentiation and development can be greatly reconstituted. Reconstitution of B-cell differentiation and development is indicative of immune system reconstitution. In general, B-cell reconstitution is compared to wild type controls. Thus, in preferred embodiments of the invention, populations of mice having inserted human variable regions possess greater than about 50% B-cell function when compared to populations of wild type mice.

Further, it is interesting to note that production of human antibodies in preference to mouse antibodies is substantially elevated in mice having a knock-out background of endogenous Ig. That is to say that mice that contain a human Ig locus and a functionally inactivated endogenous heavy chain Ig locus produce human antibodies at a rate of approximately 100 to 1000 fold as efficiently as mice that only contain a human Ig locus and are not inactivated for the endogenous locus.

### **Isotype Switching**

As is discussed in detail herein, as expected, XenoMouse II mice undergo efficient and effective isotype switching from the human transgene encoded mu isotype to the transgene encoded gamma-2 isotype. We have also developed XenoMouse II strains that contain and encode the human gamma-4 constant region. As mentioned above, mice in accordance with the invention can additionally be equipped with other human constant regions for the generation of additional isotypes. Such isotypes can include genes encoding  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\gamma_4$ ,  $\alpha$ ,  $\epsilon$ ,  $\delta$ , and other constant region encoding genes. Alternative constant regions can be included on the same transgene, i.e., downstream from the human mu constant region, or, alternatively, such other constant regions can be included on another chromosome. It will be appreciated that where such other constant regions are included on the same chromosome as the chromosome including the human mu constant region encoding transgene, cis-switching to the other isotype or isotypes can be accomplished. On the other hand, where such other constant region is included on a different chromosome from the chromosome containing the mu constant region encoding transgene, trans-switching to the other isotype or isotypes can be accomplished. Such arrangement allows tremendous flexibility in the design and construction of mice for the generation of antibodies to a wide array of antigens.

It will be appreciated that constant regions have known switch and regulatory sequences that

they are associated with. All of the murine and human constant region genes had been sequenced and published by 1989. See Honjo et al. "Constant Region Genes of the Immunoglobulin Heavy Chain and the Molecular Mechanism of Class Switching" in *Immunoglobulin Genes* (Honjo et al. eds., Academic Press (1989)), the disclosure of which is hereby incorporated by reference. For example, in U.S. Patent Application Serial No. 07/574,748, the disclosure of which is hereby incorporated by reference, the cloning of the human gamma-1 constant region was prophesized based on known sequence information from the prior art. It was set forth that in the unrearranged, unswitched gene, the entire switch region was included in a sequence beginning less than 5 kb from the 5' end of the first  $\gamma$ -1 constant exon. Therefore the switch region was also included in the 5' 5.3 kb HindIII fragment that was disclosed in Ellison et al. *Nucleic Acids Res.* **10**:4071-4079 (1982). Similarly, Takahashi et al. *Cell* **29**:671-679 (1982) also reported that the fragment disclosed in Ellison contained the switch sequence, and this fragment together with the 7.7 kb HindIII to BamHI fragment must include all of the sequences necessary for the heavy chain isotype switching transgene construction.

Thus, it will be appreciated that any human constant region of choice can be readily incorporated into mice in accordance with the invention without undue experimentation. Such constant regions can be associated with their native switch sequences (i.e., a human  $\gamma_{1, 2, 3, \text{ or } 4}$  constant region with a human  $\gamma_{1, 2, 3, \text{ or } 4}$  switch, respectively) or can be associated with other switch sequences (i.e., a human  $\gamma_4$  constant region with a human  $\gamma_2$  switch). Various 3' enhancer sequences can also be utilized, such as mouse, human, or rat, to name a few. Similarly other regulatory sequences can also be included.

As an alternative to, and/or in addition to, isotype switching *in vivo*, B-cells can be screened for secretion of "chimeric" antibodies. For example, the L6 mice, in addition to producing fully human IgM antibodies, produce antibodies having fully human heavy chain V, D, J regions coupled to mouse constant regions, such as a variety of gammas (i.e., mouse IgG1, 2, 3, 4) and the like. Such antibodies are highly useful in their own right. For example, human constant regions can be included on the antibodies through *in vitro* isotype switching techniques well known in the art. Alternatively, and/or in addition, fragments (i.e., F(ab) and F(ab')<sub>2</sub> fragments) of such antibodies can be prepared which contain little or no mouse constant regions.

As discussed above, the most critical factor to antibody production is specificity to a desired antigen or epitope on an antigen. Class of the antibody, thereafter, becomes important according to the therapeutic need. In other words, will the therapeutic index of an antibody be enhanced by providing a particular isotype or class? Consideration of that question raises issues of complement fixation and the like, which then drives the selection of the particular class or isotype of antibody. Gamma constant regions assist in affinity maturation of antibodies. However, the inclusion of a human gamma constant region on a transgene is not required to achieve such maturation. Rather, the process appears to proceed as well in connection with mouse gamma constant regions which are trans-switched onto the mu encoded transgene.

#### MATERIALS AND METHODS

The following Materials and Methods were utilized in connection with the generation and characterization of mice in accordance with the present invention. Such Materials and Methods are meant to be illustrative and are not limiting to the present invention.

Cloning Human Ig-derived YACs: The Washington University (Brownstein et al., 1989) and the CEPH (Abertsen et al., 1990) human-YAC libraries were screened for YACs containing sequences from the human heavy and kappa light chain loci as previously described (Mendez et al. 1995). Cloning and characterization of 1H and 1K YACs was described by Mendez et al., (1995). 3H and 4H YACs were identified from the Washington University library using a V<sub>H</sub>3 probe (0.55 kb PstI/NcoI, Berman et al, 1988). The 17H YAC was cloned from the GM1416 YAC library and determined to contain 130 kb of heavy chain variable sequences and a 150 kb chimeric region at its 3' end Matsuda et. al., 1993. 2K and 3K YACs were recovered from the CHEF library using V<sub>K</sub>II-specific primer (Albertsen et al., 1990).

YAC targeting and recombination: Standard methods for yeast growth, mating, sporulation, and phenotype testing were employed (Sherman et al, 1986). Targeting of YAC's and YAC vector arms with yeast and mammalian selectable markers, to facilitate the screening of YAC recombinants in yeast of YAC integration into cells, was achieved by lithium acetate transformation (Scheistl and

Geitz (1989). After every targeting or recombination step the modified YAC(s) was analyzed by pulsed field gel electrophoresis and standard Southern Blots to determine the integrity of all sequences.

YAC targeting vectors were used for the interconversion of centric and acentric arms to reorient 17H and to retrofit its 5' arm with LEU2 and URA3 genes and its 3' arm with the HIS3 gene.

5 See Fig. 1a and Mendez et al, 1993. The 4H centric arm was retrofitted with the yeast ADE2 gene and the human HPRT selectable markers. For the first recombination step, a diploid yeast strain was created and selected in which all three YACs 17H, 3H, and 4H were present, intact, and stably maintained. A three-way homologous recombination between the YAC overlapping regions was induced by sporulation and the desired recombinant was found by the selection of the outer yeast  
10 selectable markers (ADE2 and HIS3) and negative selection (loss) of the internal marker URA3. The successful recombination created a 880 kb YAC containing 80% of the IgH variable region, starting at V<sub>H</sub>2-5 and extending 20 kb 5' of the V<sub>H</sub>3-65 gene. For the recombination of the 880 kb YAC to 1H, 1H was retrofitted with pICL, which adds the LYS2 gene to the centric arm (Hermanson et al., 1991). Using standard yeast mating, a diploid strain was selected containing both 1H and the 880  
15 kb YAC. Upon sporulation and by use of overlapping homology, YAC-yeast recombination was carried out. With positive selection for the outer yeast markers (ADE2 and URA3) and screening for the loss of the internal markers (TRP1, LYS2, HIS3), an intact 970 kb YAC consisting of approximately 66 V<sub>H</sub> segments, starting at V<sub>H</sub>6-1 and ending at V<sub>H</sub>3-65 was found. The YAC also contained the major D gene clusters, J<sub>H</sub> genes, the intronic enhancer (E<sub>μ</sub>), C<sub>μ</sub>, up to 25 kb past C<sub>δ</sub>,  
20 in germline configuration. This 970 kb YAC was then retrofitted with a targeting vector including a 23 kb EcoRI genomic fragment of the human  $\gamma$ -2 gene, including its switch and regulatory elements, a 7 kb XbaI fragment of the murine heavy chain 3' enhancer, neomycin gene driven by the metallothionine promoter (MMTNeo), and the yeast LYS2 gene. This vector, while bringing in these sequences on the 3' YAC arm, disrupts the URA3 gene.

25 As a first step toward creating yK2 YAC, by standard yeast mating a diploid yeast strain was selected in which retrofitted 1K and 3K YACs were both present, intact, and stably maintained. Using the same process as described in connection with the IgH construction, YAC-yeast recombination was carried out. Through use of positive selection for the outer yeast markers (LYS2, TRP1) and the screening for the loss of internal markers (URA3, TRP1), an intact 800 kb



recombinant product was found which contained 32  $V_{\kappa}$  starting at  $V_{\kappa-B3}$  and ending at  $V_{\kappa-Op11}$ . The 800 kb YAC contains a deletion of approximately 100 kb starting at  $V_{\kappa-Lp-13}$  and ending at  $V_{\kappa-Lp-5}$ . However, the YAC is in germline configuration from  $V_{\kappa-Lp-13}$  to 100 kb past  $V_{\kappa-Op-1}$ . The YAC also contains  $J_{\kappa}$ , the intronic and 3' enhancers, the constant  $C_{\kappa}$ , and Kde.

5

**YAC introduction into ES cells and mice:** YAC-containing yeast spheroplasts were fused with E14.TG3B1 ES cells as described (Jakobovits et al., 1993a; Green et al., 1994). HAT-resistant colonies were expanded for analysis. YAC integrity was evaluated by Southern Blot analysis using protocols and probes described in Berman et al., (1988) and Mendez et al., (1994) and hybridization conditions as described in Gemmil et al., (1991). Chimeric mice were generated by microinjection of ES cells into C57BL/6 blastocysts. YAC-containing offspring were identified by PCR analysis of tail DNA as described (Green et al., 1994). YAC integrity was evaluated by Southern Blot analysis using probes and conditions previously described, except that the blot probed with human  $V_H3$  was washed at 50°C.

15

**Flow cytometry analysis:** Peripheral blood and spleen lymphocytes obtained from 8-10 week old XenoMice and control mice were purified on Lympholyte M (Accurate) and treated with purified anti-mouse CD32/CD16 Fc receptor (Pharmingen, 01241D) to block non-specific binding to Fc receptors, stained with antibodies and analyzed on a FACStar<sup>PLUS</sup> (Becton Dickinson, CELLQuest software). Antibodies used: allophycocyanin (APC) anti-B220 (Pharmingen, 01129A); biotin anti-human IgM (Pharmingen, 08072D); biotin anti-mouse IgM (Pharmingen, 02202D); fluorescein isothiocyanate (FITC) goat F(ab')<sub>2</sub> anti-human IgD (Southern Biotechnology, 2032-02); FITC anti-mouse IgD<sup>a</sup> (Pharmingen, 05064D); FITC anti-mIgD<sup>b</sup> (Pharmingen, 05074D); FITC anti-mouse  $\lambda$  (Pharmingen, 02174D); PE anti-human  $\kappa$  (Pharmingen, 08175A); PE anti-mouse  $\kappa$  (Pharmingen, 02155A.) RED613<sup>TM</sup>-streptavidin (GibcoBRL, 19541-010) was used to detect biotinylated antibodies.

20

25

**Immunization and hybridoma generation:** XenoMice (8 to 10 weeks old) were immunized intraperitoneally with 25  $\mu$ g of recombinant human IL-8 or with 5  $\mu$ g TNF- $\alpha$  (Biosource

International) emulsified in complete Freund's adjuvant for the primary immunization and in incomplete Freund's adjuvant for the additional immunizations carried out at two week intervals. For EGFR immunization, XenoMice were immunized intraperitoneally with  $2 \times 10^7$  A431 (ATCC CRL-7907) cells resuspended in phosphate buffered saline (PBS). This dose was repeated three times. Four days before fusion, the mice received a final injection of antigen or cells in PBS. Spleen and lymph node lymphocytes from immunized mice were fused with the non-secretory myeloma NSO-bcl2 line (Ray and Diamond, 1994), and were subjected to HAT selection as previously described (Galfre and Milstein, 1981).

**ELISA assay:** ELISA for determination of antigen-specific antibodies in mouse serum and in hybridoma supernatants were carried out as described (Coligan et al., 1994) using recombinant human IL-8 and TNF- $\alpha$  and affinity-purified EGFR from A431 cells (Sigma, E-3641) to capture the antibodies. The concentration of human and mouse immunoglobulins were determined using the following capture antibodies: rabbit anti-human IgG (Southern Biotechnology, 6145-01), goat anti-human Ig $\kappa$  (Vector Laboratories, AI-3060), mouse anti-human IgM (CGI/ATCC, HB-57), for human  $\gamma$ ,  $\kappa$ , and  $\mu$  Ig, respectively, and goat anti-mouse IgG (Caltag, M 30100), goat anti-mouse Ig $\kappa$  (Southern Biotechnology, 1050-01), goat anti-mouse IgM (Southern Biotechnology, 1020-01), and goat anti-mouse  $\lambda$  (Southern Biotechnology, 1060-01) to capture mouse  $\gamma$ ,  $\kappa$ ,  $\mu$ , and  $\lambda$  Ig, respectively. The detection antibodies used in ELISA experiments were goat anti-mouse IgG-HRP (Caltag, M-30107), goat anti-mouse Ig $\kappa$ -HRP (Caltag, M 33007), mouse anti-human IgG2-HRP (Southern Biotechnology, 9070-05), mouse anti-human IgM-HRP (Southern Biotechnology, 9020-05), and goat anti-human kappa-biotin (Vector, BA-3060). Standards used for quantitation of human and mouse Ig were: human IgG $_2$  (Calbiochem, 400122), human IgM $\kappa$  (Cappel, 13000), human IgG $_2$  $\kappa$  (Calbiochem, 400122), mouse IgG $\kappa$  (Cappel 55939), mouse IgM $\kappa$  (Sigma, M-3795), and mouse IgG $_3$  $\lambda$  (Sigma, M-9019).

**Determination of affinity constants of fully human Mabs by BIAcore:** Affinity measurement of purified human monoclonal antibodies, Fab fragments, or hybridoma supernatants by plasmon resonance was carried out using the BIAcore 2000 instrument, using general procedures

outlined by the manufacturers.

Kinetic analysis of the antibodies was carried out using antigens immobilized onto the sensor surface at a low density: human IL-8 -81 RU, soluble EGFR purified from A431 cell membranes (Sigma, E-3641)- 303 RU, and TNF- $\alpha$ - 107 RU (1,000 RU correspond to about 1 ng/mm<sup>2</sup> of immobilized protein). The dissociation (kd) and association (ka) rates were determined using the software provided by the manufacturers, BIAevaluation 2.1.

**Affinity measurement by radioimmunoassay:** <sup>125</sup>I-labeled human IL-8 ( $1.5 \times 10^{-11}$  M or  $3 \times 10^{-11}$  M) was incubated with purified anti-IL-8 human antibodies at varying concentrations ( $5 \times 10^{-13}$  M to  $4 \times 10^{-9}$  M) in 200  $\mu$ l of PBS with 0.5% BSA. After 15 hrs. incubation at room temperature, 20  $\mu$ l of Protein A Sepharose CL-4B in PBS (1/1, v/v) was added to precipitate the antibody-antigen complex. After 2 hrs. incubation at 4°C, the antibody-<sup>125</sup>I-IL-8 complex bound to Protein A Sepharose was separated from free <sup>125</sup>I-IL-8 by filtration using 96-well filtration plates (Millipore, Cat. No. MADVN65), collected into scintillation vials and counted. The concentration of bound and free antibodies was calculated and the binding affinity of the antibodies to the specific antigen was obtained using Scatchart analysis (2).

**Receptor binding assays:** The IL-8 receptor binding assay was carried out with human neutrophils prepared either from freshly drawn blood or from buffy coats as described (Lustig-Marasimhan et al., 1995). Varying concentrations of antibodies were incubated with 0.23 nM [<sup>125</sup>I]IL-8 (Amersham, IM-249) for 30 min at 4 °C in 96-well Multiscreen filter plates (Millipore, MADV N6550) pretreated with PBS binding buffer containing 0.1% bovine serum albumin and 0.02% NaN<sub>3</sub> at 25°C for 2 hours.  $4 \times 10^5$  neutrophils were added to each well, and the plates were incubated for 90 min at 4°C. Cells were washed 5 times with 200  $\mu$ l of ice-cold PBS, which was removed by aspiration. The filters were air-dried, added to scintillation fluid, and counted in a scintillation counter. The percentage of specifically bound [<sup>125</sup>I]IL-8 was calculated as the mean cpm detected in the presence of antibody divided by cpm detected in the presence of buffer only.

Binding assays for TNF receptor were performed in a similar manner as the IL-8 assays described above. However, the human monocyte line U937 was utilized instead of the neutrophil line

used in connection with the IL-8 assays. Antibodies were preincubated with 0.25 nM [<sup>125</sup>I]TNF (Amersham, IM-206). 6 x 10<sup>5</sup> U937 cells were placed in each well.

The EGF receptor binding assay was carried out with A431 cells (0.4 x 10<sup>6</sup> cells per well) which were incubated with varying concentrations of antibodies in PBS binding buffer for 30 minutes at 4°C. 0.1 nM [<sup>125</sup>I]EGF (Amersham, IM-196) was added to each well, and the plates were incubated for 90 min at 4°C. The plates were washed five times, air-dried and counted in a scintillation counter. Anti-EGFR mouse antibodies 225 and 528 (Calbiochem) were used as controls.

**Repertoire analysis of human Ig transcripts expressed in XenoMice and their derived**

**human Mabs:** Poly(A)<sup>+</sup> mRNA was isolated from spleen and lymph nodes of unimmunized and immunized XenoMice using a Fast-Track kit (Invitrogen). The generation of random primed cDNA was followed by PCR. Human V<sub>H</sub> or human V<sub>κ</sub> family specific variable region primers (Marks et. al., 1991) or a universal human V<sub>H</sub> primer, MG-30 (CAGGTGCAGCTGGAGCAGTCIGG) was used in conjunction with primers specific for the human C<sub>μ</sub> (hμP2) or C<sub>κ</sub> (hκP2) constant regions as previously described (Green et al., 1994), or the human γ2 constant region MG-40d; 5'-GCTGAGGGAGTAGAGTCCTGAGGA-3'. PCR products were cloned into pCRII using a TA cloning kit (Invitrogen) and both strands were sequenced using Prism dye-terminator sequencing kits and an ABI 377 sequencing machine. Sequences of human Mabs-derived heavy and kappa chain transcripts were obtained by direct sequencing of PCR products generated from poly(A<sup>+</sup>) RNA using the primers described above. All sequences were analyzed by alignments to the "V BASE sequence directory" (Tomlinson et al., MRC Centre for Protein Engineering, Cambridge, UK) using MacVector and Geneworks software programs.

**Preparation and purification of antibody Fab fragments:** Antibody Fab fragments were produced by using immobilized papain (Pierce). The Fab fragments were purified with a two step chromatographic scheme: HiTrap (Bio-Rad) Protein A column to capture Fc fragments and any undigested antibody, followed by elution of the Fab fragments retained in the flow-through on strong cation exchange column (PerSeptive Biosystems), with a linear salt gradient to 0.5 M NaCl. Fab fragments were characterized by SDS-PAGE and MALDI-TOF MS under reducing and non-reducing

conditions, demonstrating the expected ~50 kD unreduced fragment and ~25 kDa reduced doublet. This result demonstrates the intact light chain and the cleaved heavy chain. MS under reducing conditions permitted the unambiguous identification of both the light and cleaved heavy chains since the light chain mass can be precisely determined by reducing the whole undigested antibody.

5

### EXAMPLES

The following examples, including the experiments conducted and results achieved are provided for illustrative purposes only and are not to be construed as limiting upon the present invention.

10

#### Example 1: Reconstruction of human heavy chain loci on YACs

In accordance with the present invention, the strategy that we utilized to reconstruct the human heavy chain and human kappa light chain variable regions was to, first, screen human-YAC libraries for YACs that spanned the large (megabase-sized) human Ig loci and, second, to recombine YACs spanning such regions into single YACs containing the desired loci predominantly in germline configuration.

The above, stepwise, YAC recombination scheme exploited the high frequency of meiotic-induced homologous recombination in yeast and the ability to select the desired recombinants by the yeast markers present on the vector arms of the recombined YACs (*See* Figure 1, and Green et al., *supra.*; *see also* Silverman et al., 1990 and denDunnen et al., 1992).

In connection with our strategy, we identified four YACs, 1H (240 kb), 2H (270 kb), 3H (300 kb), and 4H (340 kb), which spanned about 830 kb, out of the about 1000 kb, of the human heavy chain variable region on chromosome 14q. YACs 1H, 2H, 3H, and 4H were used for reconstruction of the locus (*See* Figure 1A). Pulsed Field Gel Electrophoresis (PFGE) and Southern blot analysis confirmed the YACs to be in intact, germline configuration, with the exception of 150 kb at the 3' end of YAC 2H which contained certain non-IgH sequences (*See* Figure 1; Matsuda et al., 1990). YAC 1H, the YAC that was previously introduced into our first generation XenoMouse™ (Green et al., *supra.*; Mendez et al., 1995), is comprised of the human C<sub>δ</sub>, C<sub>μ</sub>, J<sub>H</sub>, and D<sub>H</sub> regions and the first 5 V<sub>H</sub> genes in germline configuration. The other three YACs cover the majority of the V<sub>H</sub> region, from

V<sub>H</sub>2-5 to V<sub>H</sub>3-65, thus contributing approximately an additional 61 different V<sub>H</sub> genes. Prior to recombination, YAC 4H was retrofitted with an HPRT selectable marker. Through utilization of the overlapping sequences contained on the YACs, the four YACs (1H, 2H, 3H, and 4H) were recombined in yeast by a stepwise recombination strategy (See Figure 1A). Such recombination strategy generated a 980 kb recombinant YAC (See Figure 1). Analysis of the YAC by PFGE and Southern blot analysis confirmed the presence of the human heavy chain locus from the C<sub>δ</sub> region to 20 kb 5' of the V<sub>H</sub>3-65 gene in germline configuration. No apparent deletions or rearrangements were observed.

The YAC acentric arm was targeted with a vector bearing the complete human  $\gamma$ 2 constant region, mouse 3' enhancer, and the neomycin resistance gene, to yield the final 1020 kb heavy chain YAC, yH2. YAC yH2 contained the majority of the human variable region i.e., 66 out of the 82 V<sub>H</sub> genes, complete D<sub>H</sub> (32 genes), and J<sub>H</sub> (6 genes) regions and three different constant regions (C $\mu$ , C $\delta$ , and C $\gamma$ ) with their corresponding regulatory sequences (See Figure 1A). This was the heavy chain construct utilized for the production of our XenoMouse II strains.

#### **Example 2: Reconstruction of human kappa light chain loci on YACs**

A similar stepwise recombination strategy was utilized for reconstruction of the human kappa light chain locus. Three YACs were identified that spanned the human kappa loci. The YACs were designated 1K, 2K and 3K. YAC 1K, which had a length of approximately 180 kb, had previously been introduced into our first generation XenoMouse™. Such YAC contained the kappa deleting element, (Kde), the kappa 3' and intronic enhancers, C <sub>$\kappa$</sub> , J <sub>$\kappa$</sub> , and the three V <sub>$\kappa$</sub>  genes on the B cluster (Green et al., 1994; Mendez et al., 1995). YAC 2K (approximately 480 kb), and 3K (approximately 380 kb) together encompass most of the kappa chain proximal variable region on chromosome 2p. A deletion of approximately 100 kb spans the L13-L5 region (Fig. 1B; Huber et al., 1993). Inasmuch as the kappa distal region duplicates the proximal region, and as the proximal V <sub>$\kappa$</sub>  genes are the ones most commonly utilized humans (Weichold et al., 1993; Cox et al., 1994), the proximal region was the focus of our reconstruction strategy (Fig. 1B). Through homologous recombination of the three YACs, an 800 kb recombinant YAC, yK2, was recovered. The size and integrity of the recombinant YAC was confirmed by PFGE and Southern blot analysis. Such analysis demonstrated that it covered

the proximal part of the human kappa chain locus, with 32  $V_{\kappa}$  genes in germline configuration except for the described deletion in the Lp region (Fig. 1B). yK2 centric and acentric arms were modified to contain the HPRT and neomycin selectable markers, respectively, as described (Materials and Methods). This was the kappa light chain construct utilized for the production of our XenoMouse II strains.

The YACs described herein, yH2 and yK2, represent the first megabase-sized reconstructed human Ig loci to contain the majority of the human antibody repertoire, predominantly in germline configuration. This accomplishment further confirmed homologous recombination in yeast as a powerful approach for successful reconstruction of large, complex, and unstable loci. The selection of stable YAC recombinants containing large portions of the Ig loci in yeast provided us with the human Ig fragments required to equip the mice with the human antibody repertoire, constant regions, and regulatory elements needed to reproduce human antibody response in mice.

**Example 3: Introduction of yH2 and yK2 YACs into ES cells**

In accordance with our strategy, we introduced the YACs, yH2 and yK2, into mouse embryonic stem (ES) cells. Once ES cells containing the YAC DNA were isolated, such ES cells were utilized for the generation of mice through appropriate breeding.

In this experiment, therefore, YACs yH2 and yK2, were introduced into ES cells via fusion of YAC-containing yeast spheroplasts with HPRT-deficient E14.TG3B1 mouse ES cells as previously described (Jakobovits et al., 1993a; Green et al., 1994). HPRT-positive ES cell clones were selected at a frequency of 1 clone/15-20x10<sup>6</sup> fused cells and were analyzed for YAC integrity by Southern and CHEF blot analyses (Fig. 2A).

Seven of thirty-five ES cell clones (referred to as L10, J9.2, L17, L18, J17, L22, L23) derived from ES cell fusion with yH2-containing yeast were found to contain all expected EcoRI and BamHI yH2 fragments detected by probes spanning the entire insert: mouse 3' enhancer, human intronic enhancer, human  $C_{\gamma}2$ ,  $C_{\delta}$ , and  $C_{\mu}$  constant regions,  $D_H$ ,  $J_H$  and all the different  $V_H$  families:  $V_H1$ ,  $V_H2$ ,  $V_H3$ ,  $V_H4$ ,  $V_H5$ , and  $V_H6$  (data shown for 5 clones in Fig. 2A). CHEF analysis further confirmed that these clones, which represent 20% of all clones analyzed, contain the entire intact yH2 YAC with no apparent deletions or rearrangements (data not shown).

ES cell clones derived from the fusion of yK2-containing yeast were similarly analyzed for YAC integrity, using probes specific for the human Kde, kappa 3' and intronic enhancers, C<sub>κ</sub>, J<sub>κ</sub>, and all of the different V<sub>κ</sub> families: V<sub>κ</sub>I, V<sub>κ</sub>II, V<sub>κ</sub>III, V<sub>κ</sub>IV, V<sub>κ</sub>VI. Twenty clones of the sixty clones had intact and unaltered YAC, which represent 30% of total clones analyzed (data shown for two ES clones in: Fig. 3A). Varying amounts of yeast genomic sequences were detected in yH2 and yK2-ES cell clones (data not shown).

These results are the first demonstration of introduction of megabase-sized constructs encompassing reconstructed human loci, predominantly in germline configuration, into mammalian cells. The relatively high frequency of intact YACs integrated into the mouse genome further validated the ES cell-yeast spheroplast fusion methodology as an effective approach for faithful introduction of large human genomic fragments into ES cells.

**Example 4: Generation of XenoMouse II strains**

In order to generate mice from the YAC DNA containing ES cells, microinjection of blastocysts was conducted, followed by breeding. Thus, yH2- and yK2-bearing ES cell clones were expanded and microinjected into mouse C57BL/6J blastocysts (Green et al., 1994) and the chimeric males produced were evaluated for germline transmission. Offspring with transmitted YAC were identified by PCR analysis and the YAC integrity was confirmed by Southern blot analysis. In all transgenic mice analyzed the YAC was shown to be in intact form (Fig.2B, 3B). All seven microinjected yH2-ES clones and two out of eight yK2-ES clones were transmitted through the mouse germline.

In order to generate mice that produced human antibodies to the exclusion of endogenous antibodies, yH2- or yK2-transgenic mice were bred with double-inactivated (DI) mouse strains. The DI mouse strains are homozygous for gene targeted-inactivated mouse heavy and kappa chain loci and thus are deficient in antibody production (Jakobovits et al., 1993b; Green et al., 1994). Two of the yH2- transgenic mouse strains L10 and J9.2, and one of the yK2-transgenic mouse strains, J23.1, were bred with DI mice to generate mice bearing YACs on an homozygous inactivated mouse heavy and kappa chain background (yH2;DI, and yK2;DI). Each of the yH2;DI transgenic strains were bred with the yK2;DI transgenic strain to generate two XenoMouse II strains, 2A-1 (L10;J23.1;DI) and



2A-2 (J9.2;J23.1;DI), respectively, containing both heavy and light chain YACs on homozygous DI background. L10 is fully homozygous and J9.2 and J23.1 are in the process of being successfully bred to homozygosity.

The integrity of the human heavy and kappa chain YACs in XenoMouse II strains was confirmed by Southern blot analysis. As shown in Fig. 2 and Fig. 3, in both XenoMouse strains analyzed, yH2 and yK2 were transmitted unaltered through multiple generations with no apparent deletions or rearrangements.

**Example 5: B-cell development and human antibody production by XenoMouse II mice**

In order to further characterize the XenoMouse II strains, we studied their B-cell development and their production of human antibodies. Reconstitution of B-cell development and antibody production in XenoMouse II strains by yH2 and yK2 YACs was evaluated by flow cytometry and ELISA. In contrast to DI mice, which completely lack mature B-cells, XenoMouse II manifested essentially normal B-cell development with the mature B-cell population in the blood totaling over 50% of the level seen in wild type mice (Fig. 4A). All B-cells were shown to express human IgM and high levels of B220 (human IgM<sup>+</sup>/B220<sup>hi</sup>), with 60% of this population also expressing human IgD. Similar results were obtained from analysis of XenoMouse spleen and lymph nodes (not shown). These results correlate well with the characteristics of mature B-cells in wild type mice, indicating proper B-cell maturation in XenoMouse.

The majority of XenoMouse B-cells (75-80%) expressed exclusively human kappa ( $\kappa$ ) light chain, whereas only about 15% expressed mouse lambda ( $\lambda$ ) light chain (Fig. 4). This light chain distribution ratio ( $h\kappa/m\lambda$ : 75:15) is comparable to that observed in wild type mice, indicating a mouse-like regulation of light chain utilization. In contrast, XenoMouse I, as described in Green et al., 1994, showed a ratio of  $h\kappa/m\lambda$ : 55:45 (data not shown). Similar observations were made for B-cells from spleen (Fig. 4B) and lymph nodes (not shown), indicating that most of XenoMouse II's B-cells produced exclusively fully human antibodies. Levels of  $m\lambda$ -expressing B-cells were reduced from 15% to 7% in XenoMouse II strains homozygous for yK2 (data not shown).

**Example 6    Generation of L6 Strain**

The L6 strain of mice were generated identically to the process described above in connection with the generation of the XenoMouse II strains. However, owing to a deletion event during the generation of the L6 ES cell line, the ES cell line, and, subsequently, the L6 mouse evolved without a portion of the sequence distal to C $\delta$ , thus, eliminating the C $\gamma$  constant region and its regulatory sequences. Following completion of breeding, the L6 mice will contain the entire yK2 construct and the entire yH2 construct, except for the missing C $\gamma$  constant region.

**Example 7:    Human Antibody Production**

Expression of human C $\mu$ , C $\gamma$ 2, and  $\kappa$  light chains were detected in unimmunized XenoMouse II sera at maximal levels of 700, 600, and 800  $\mu$ g/ml, respectively. To determine how these values compared to wild-type, we measured maximal levels of mouse C $\mu$ , C $\gamma$ 2, and  $\kappa$  light chains in C57BL/6J x 129 mice kept under similar pathogen-free conditions. The values for C $\mu$ , C $\gamma$ 2, and  $\kappa$  light chain in wild-type mice were 400, 2000, and 2000  $\mu$ g/ml, respectively. Upon immunization, the human  $\gamma$  chain levels increased to approximately 2.5 mg/ml. The concentration of mouse  $\lambda$  was only 70  $\mu$ g/ml, further confirming the preferential use of human kappa chain.

These findings confirmed the ability of the introduced human Ig YACs to induce proper Ig gene rearrangement and class switching and to generate significant levels of fully human IgM and IgG antibodies before and after immunization.

**Example 8:    A diverse human antibody repertoire in XenoMouse II**

In order to further understand the reconstitution of the antibody repertoire in XenoMouse II strains, we challenged mice with several antigens, and prepared hybridoma cell lines secreting such antibodies. As will be understood, recapitulation of the human antibody response in mice requires diverse utilization of the different human variable genes contained on yH2 and yK2 YACs. The diversity of the human antibodies generated by XenoMouse II strains was determined by cloning and sequencing human heavy chain ( $\mu$  and  $\gamma$ ) and kappa light chain transcripts from XenoMouse lymph nodes. Based upon our data to date, sequence analysis demonstrates that XenoMouse II utilizes at

least 11 out of the 37 functional  $V_H$  genes present on yH2, eight different  $D_H$  segments and three  $J_H$  genes ( $J_{H3}$ ,  $J_{H4}$ ,  $J_{H6}$ ) (Table III;  $J_{H5}$  was also detected in connection with our sequencing antibodies from hybridomas). V-D-J sequences were linked to human  $\mu$  or  $\gamma 2$  constant regions (not shown).

The  $V_H$  genes utilized are widely distributed over the entire variable region and represent four out of the seven  $V_H$  families (Table III). The predominant utilization of V genes from  $V_{H3}$  and  $V_{H4}$  families is similar to the  $V_H$  usage pattern in adult humans, which is proportional to family size (Yamada et al. 1991; Brezinshek et al., 1995). The predominant usage of  $J_{H4}$  is also reminiscent of that detected in human B-cells (Brezinshek et al., 1995). Addition of non-germline nucleotides (N-additions) at both V-D and D-J joinings, ranging from 1-12 bp, were also observed. Such N-additions produced complementary determining regions 3 (CDR3s) with lengths of from 8 to about 19 amino acid residues, which is very comparable to that observed in adults human B-cells (Yamada et al. 1991; Brezinshek et al., 1995). Such CDR3 lengths observed in the XenoMouse II are much longer than CDR3 lengths ordinarily observed in mice (Feeny, 1990).

A highly diverse repertoire was also found in the ten kappa chain transcripts sequenced. In addition to displaying 8 out of the 25  $V_\kappa$  functional open reading frames (ORFs) present on yK2, all of the  $J_\kappa$  genes were detectable (Table IV). The different  $V_\kappa$  genes utilized were widely dispersed throughout yK2, representing all four major  $V_\kappa$  gene families. All  $V_\kappa J_\kappa$  recombination products were linked properly to  $C_\kappa$  sequences. The paucity of N-additions in our transcripts is in agreement with the greatly reduced terminal deoxynucleotide transferase activity at the stage of kappa chain rearrangement. The average CDR3 length of 9-10 amino acids that we observed in the kappa chain transcripts is identical to that observed in human B-cells (Marks et al., 1991).

In Tables III and IV below, repertoire analyses of human heavy and kappa light chain transcripts expressed in XenoMouse II strains are presented. Human  $\mu$ ,  $\gamma$ , and  $\kappa$  specific mRNAs were amplified by PCR, cloned and analyzed by sequencing as described in Materials and Methods. Table III shows a series of nucleotide sequences of 12 unique human heavy chain clones, divided into  $V_H$ , D,  $J_H$  and N segments, as identified by homology with published germline sequences (Materials and Methods). Each D segment assignment is based on at least 8 bases of homology. Table IV shows a series of nucleotide sequences of V-J junctions of 8 independent human  $\kappa$  clones. The sequences are divided into  $V_\kappa$ ,  $J_\kappa$  and N segments and identified based on homology to published  $V_\kappa$

and J<sub>κ</sub> sequences. In each of the Tables N-additions and deletions (indicated as   ) were determined by their lack of sequence homology to V, D, or J sequences.

**TABLE III**  
**Repertoire Analysis of Human Heavy Chain Transcripts**

Clone	V <sub>H</sub>	N	D <sub>H</sub>	N	J <sub>H</sub>
A2.2.1	5-51 (DP73) TTACTGTGCGAGACA	4 (TAGG)	XP5rc AATCAT	12 (GGGAGCTACGGG)	JH4 ____ GACTACTGGGGC
B2.1.5	3-33 (DP-50) TTACTGTGCGAGAGA	7 (TCGGGGA)	3rc AATAGCA	7 (CTGGCCT)	JH4 __CTTTGACTACTGGGGC
B4.2.4	3-15 (DP-38) TTACTGTACCACAGA	1 (G)	K1 GGCTAC	11 (ACTAACTACCC)	JH6 ____CTACTACTACTACGGT
B4.2.5	4-59 (DP-71) TTACTGTGCGAGAGA	10 (TAGGAGTGTT)	4 GTAGTACCAGCTGCTAT	6 (ACCCAA)	JH6 __ACTACTACTACTACGGT
D2.2.5	4-34 (DP-63) TTACTGTGCGAGAG_	2 (GG)	N1 GCAGCAGCTG	4 (CCCT)	JH4 __CTTTGACTACTGGGGC
D2.1.3	3-48 (DP-51) TTACTGTGCGAGAGA	4 (TCTT)	XP1 GATATTTTGACTGGT	2 (CT)	JH6 ____CTACTACTACTACGGT
D2.1.3	4-31 (DP-65) TTACTGTGCGAGAGA	2 (GA)	A4 GACTGCAG	5 (CGGTT)	JH4 ____TTTGACTACTGGGGC
A2.2.4	3-21 (DP-77) TTACTGTGCGAGAGA	2 (TT)	IR3 GGGGCTGG	3 (ACC)	JH6 _TACTACTACTACTACGGT
D4.2.11	4-4/4.35 ATTACTGTGCGA	1 (A)	N1 TATAGCAGTGGCTGGT	2 (GT)	JH4 __CTTTGACTACTGGGGC
C1.2.1	1-18 (DP-14) TATTACTGTGCGAG_	0	XP'1/21-7 GTTA	0	JH4 _____GACTACTGGGGC
C3.1.2	4-39 (DP-79) TATTACTGTGCG__	3 (GCC)	2 GGATATAGTAGTGG	6 (TCGGGC)	JH4 __CTTTGACTACTGGGGC
D2.2.7	5-51 (DP73) TTACTGTGCGAGACA	4 (TGGC)	K1 AGTGGCT	9 (GGTACTCTG)	JH3 ATGCTTTGATATCTGGGG

**TABLE IV**  
**Repertoire Analysis of Human Kappa Light Chain Transcripts**

Clone	V <sub>κ</sub>	N	J <sub>κ</sub>
F2.2.3	02 (DPK9) TTAAACGAACAGTACCCC__	0	Jκ5 GATCACCTTCGGCCAA
F4.1.8	L5 (DPK5) ACAGGCTAACAGTTCCCTC_	0	Jκ1 __GGACGTTCGGCCAA
F4.1.6	A20 (DPK4) AAGTATAACAGTGCCCC__	0	Jκ3 ATTCACTTTCGGCCCT
F2.2.5	08 ACAGTATGATAATCTCCC__	0	Jκ4 GCTCACTTTCGGCGGA
F2.1.5	L1 AAAGTATAATAGTTACCC__	0	Jκ5 GATCACCTTCGGCCAA
F2.1.4	A30 CAGCATAATAGTTACCC__	0	Jκ3 ATTCACTTTCGGCCCT
F2.1.3	B3 (DPK24) AATATTATAGTACTCC__	0	Jκ4 GCTCACTTTCGGCGGA
F4.1.3	A27 (DPK22) CAGTATGGTAGCTCACCTC	1 (G)	Jκ2 __CACTTTTGCCAG

These results, together with sequences of XenoMouse-derived hybridomas described later, demonstrate a highly diverse, adult human-like utilization of V, D, and J genes, which appears to demonstrate that the entire human heavy and kappa chain variable regions present on the yH2 and the yK2 YACs are accessible to the mouse system for antibody rearrangement and are being utilized in a non-position-biased manner. In addition, the average length of N-additions and CDR3s for both the heavy and kappa chain transcripts, is very similar to that seen in adult human B-cells, indicating that the YAC DNA contained in the mice direct the mouse machinery to produce an adult human-like immune repertoire in mice.

In connection with the following Examples, we prepared high affinity antibodies to several antigens. In particular, antigens were prepared to human IL-8 and human EGFR. The rationale for the selection of IL-8 and EGFR is as follows.

IL-8 is a member of the C-X-C chemokine family. IL-8 acts as the primary chemoattractant for neutrophils implicated in many diseases, including ARDS, rheumatoid arthritis, inflammatory bowel disease, glomerulonephritis, psoriasis, alcoholic hepatitis, reperfusion injury, to name a few.

Moreover, IL-8 is a potent angiogenic factor for endothelial cells. In Figures 22-28, we demonstrate that human anti-IL-8 antibodies derived from XenoMouse II strains are effective in inhibiting IL-8's actions in a number of pathways. For example, Figure 22 shows blockage of IL-8 binding to human neutrophils by human anti-IL-8. Figure 23 shows inhibition of CD11b expression on human neutrophils by human anti-IL-8. Figure 24 shows inhibition of IL-8 induced calcium influx by human anti-IL-8 antibodies. Figure 25 shows inhibition of IL-8 RB/293 chemotaxis by human anti-IL-8 antibodies. Figure 26 is a schematic diagram of a rabbit model of human IL-8 induced skin inflammation. Figure 27 shows the inhibition of human IL-8 induced skin inflammation in the rabbit model of Figure 26 with human anti-IL-8 antibodies. Figure 28 shows inhibition of angiogenesis of endothelial cells on a rat corneal pocket model by human anti-IL-8 antibodies.

EGFR is viewed as an anti-cancer target. For example, EGFR is overexpressed, up to 100 fold, on a variety of cancer cells. Ligand (EGF and TNF) mediated growth stimulation plays a critical role in the initiation and progression of certain tumors. In this regard, EGFR antibodies inhibit ligand binding and lead to the arrest of tumor cell growth, and, in conjunction with chemotherapeutic agents, induces apoptosis. Indeed, it has been demonstrated that a combination of EGFR Mabs resulted in tumor eradication in murine xenogeneic tumor models. Imclone has conducted Phase I clinical utilizing a chimeric Mab (C225) that proved to be safe. In Figures 31-33, we demonstrate data related to our human anti-EGFR antibodies. Figure 30 shows heavy chain amino acid sequences of human anti-EGFR antibodies derived from XenoMouse II strains. Figure 31 shows blockage EGF binding to A431 cells by human anti-EGFR antibodies. Figure 32 shows inhibition of EGF binding to SW948 cells by human anti-EGFR antibodies. Figure 33 shows that human anti-EGFR antibodies derived from XenoMouse II strains inhibit growth of SW948 cells *in vitro*.

**Example 9: High affinity, antigen-specific human Mabs produced by XenoMouse II**

We next asked whether the demonstrated utilization of the large human repertoire in XenoMouse II could be harnessed to generate human antibodies to multiple antigens, in particular, human antigens of significant clinical interest.

Accordingly, individual XenoMouse II pups were challenged each with one of three different antigen targets, human IL-8, human EGFR and human TNF- $\alpha$ . Antigens were administered in two

different forms, either as soluble protein, in the case of IL-8 and TNF- $\alpha$  or expressed on the surface of cells (A431 cells), in the case of EGFR. For all three antigens, ELISAs performed on sera from immunized mice indicated a strong antigen-specific human antibody (IgG, Ig $\kappa$ ) response with titers as high as  $1.3 \times 10^6$ . Negligible mouse  $\lambda$  response was detected.

5 Hybridomas were derived from spleen or lymph node tissues by standard hybridoma technology and were screened for secretion of antigen-specific human Mabs by ELISA.

An IL-8 immunized XenoMouse II yielded a panel of 12 hybridomas, all secreting fully human (hIgG $_2\kappa$ ) Mabs specific to human IL-8. Antibodies from four of these hybridomas, D1.1, K2.2, K4.2, and K4.3, were purified from ascitic fluid and evaluated for their affinity for human IL-8 and their  
10 potency in blocking binding of IL-8 to its receptors on human neutrophils.

Affinity measurements were performed by solid phase measurements of both whole antibody and Fab fragments using surface plasmon resonance in BIAcore and in solution by radioimmunoassay (Materials and Methods). As shown in Table V, affinity values measured for the four Mabs ranged from  $1.1 \times 10^9$  to  $4.8 \times 10^{10}$  M $^{-1}$ . While there was some variation in the techniques employed, affinity  
15 values for all four antibodies were consistently higher than  $10^9$  M $^{-1}$ .

ELISA analysis confirmed that these four antibodies were specific to human IL-8 and did not cross-react with the closely related chemokines MIP-1 $\alpha$ , GRO $\alpha$ ,  $\beta$ , and  $\gamma$ , ENA-78, MCP-1, or RANTES (data not shown). Further, competition analysis on the BIAcore indicated that the antibodies recognize at least two different epitopes (data not shown). All antibodies inhibit IL-8  
20 binding to human neutrophils as effectively as the murine anti-human IL-8 neutralizing antibody, whereas a control human IgG $_2\kappa$  antibody did not (Fig. 5A).

Fusion experiments with EGFR-immunized Xenomouse II yielded a panel of 25 hybridomas, all secreting EGFR-specific human IgG $_2\kappa$  Mabs. Of the thirteen human Mabs analyzed, four (E2.1, E2.4, E2.5, E2.11) were selected for their ability to compete with EGFR-specific mouse antibody  
25 225, which has previously been shown to inhibit EGF-mediated cell proliferation and tumor formation in mice (Sato et al., 1983). These human antibodies, purified from ascitic fluid, were evaluated for their affinity for EGFR and neutralization of EGF binding to cells. The affinities of these antibodies for EGFR, as determined by BIAcore measurements, ranged from  $2.9 \times 10^9$  to  $2.9 \times 10^{10}$  M $^{-1}$  (Table V).

All four anti-EGFR antibodies completely blocked EGF binding to A431 cells (Fig. 5B),

demonstrating their ability to neutralize its binding to both high and low affinity receptors on these cells (Kawamoto et al., 1983). Complete inhibition of EGF binding to EGFR expressed on human SW948 human lung carcinoma cells by all four anti-EGFR human antibodies was also observed (data not shown). In both cases, the fully human antibodies were as effective in inhibition of EGF binding as the anti-EGFR mouse antibody 225 and more potent than the 528 antibody (Gill et al., 1983). In both cell assays, a control human IgG<sub>2</sub>κ antibody did not affect EGF binding (Fig. 5B and data not shown).

Fusion experiments with TNF-α immunized Xenomouse II yielded a panel of 12 human IgG<sub>2</sub>κ antibodies. Four out of the 12 were selected for their ability to block the binding of TNF-α to its receptors on U937 cells (Fig. 5C). The affinities of these antibodies were determined to be in the range of 1.2-3.9x10<sup>9</sup> M<sup>-1</sup> (Table V).

The described Xenomouse-derived hybridomas produced antibodies at concentrations in the range of 2-19 µg/ml in static culture conditions. Characterization of the purified antibodies on protein gels under non-reducing conditions revealed the expected apparent molecular weight of 150 kD for the IgG<sub>2</sub>κ antibody. Under reducing conditions the expected apparent molecular weights of 50 kD for the heavy and 25 kD for the light chain were detected (data not shown).

Table V, below, shows affinity constants of XenoMouse-derived antigen-specific fully human Mabs. The affinity constants of XenoMouse-derived human IgG<sub>2</sub>κ Mabs specific to IL-8, EGFR, and TNF-α were determined by BIAcore or by radioimmunoassay as described in Materials and Methods. The values shown for IL-8 and EGFR are representative of independent experiments carried out with purified antibodies, while the values shown for TNF-α are from experiments carried out with hybridoma supernatants.



**TABLE V**

Human Mab IgG, $\kappa$	Antigen	$k_a$ ( $M^{-1}S^{-1}$ )	$k_d$ ( $S^{-1}$ )	$KA$ ( $M^{-1}$ )	$KD$ (M)	Surface Density [RU]	Radio Immunoassay ( $M^{-1}$ )
Solid Phase Measurements						Solution	
D1.1	IL-8	$2.7 \times 10^6$	$9.9 \times 10^{-4}$	$2.7 \times 10^9$	$3.7 \times 10^{-10}$	81	$2.0 \times 10^{10}$
D1.1 Fab	IL-8	$2.1 \times 10^6$	$2.1 \times 10^{-3}$	$1.1 \times 10^9$	$8.8 \times 10^{-10}$	81	$4.9 \times 10^{11}$
K2.2	IL-8	$0.9 \times 10^6$	$2.3 \times 10^{-4}$	$4.0 \times 10^9$	$2.5 \times 10^{-10}$	81	$1.0 \times 10^{10}$
K4.2	IL-8	$2.5 \times 10^6$	$4.1 \times 10^{-4}$	$6.3 \times 10^9$	$1.6 \times 10^{-10}$	81	ND
K4.3	IL-8	$4.3 \times 10^6$	$9.4 \times 10^{-4}$	$4.5 \times 10^9$	$2.2 \times 10^{-10}$	81	$2.1 \times 10^{11}$
K4.3 Fab	IL-8	$6.0 \times 10^6$	$2.1 \times 10^{-3}$	$2.9 \times 10^9$	$3.4 \times 10^{-10}$	81	
						ELISA (M)	
E1.1	EGFR	$1.9 \times 10^6$	$6.5 \times 10^{-4}$	$2.9 \times 10^9$	$3.46 \times 10^{-10}$	303	$1.1 \times 10^{-10}$
E2.5	EGFR	$2.1 \times 10^6$	$1.8 \times 10^{-4}$	$1.2 \times 10^{10}$	$8.44 \times 10^{-11}$	303	$3.6 \times 10^{-10}$
E2.11	EGFR	$1.7 \times 10^6$	$4.7 \times 10^{-4}$	$3.7 \times 10^9$	$2.68 \times 10^{-10}$	303	$1.1 \times 10^{-10}$
E2.4	EGFR	$2.8 \times 10^6$	$9.78 \times 10^{-5}$	$2.9 \times 10^{10}$	$3.5 \times 10^{-11}$	818	$1.1 \times 10^{-10}$
T22.1	TNF- $\alpha$	$1.6 \times 10^6$	$1.3 \times 10^{-3}$	$1.2 \times 10^9$	$8.06 \times 10^{-10}$	107	
T22.4	TNF- $\alpha$	$2.4 \times 10^6$	$4.6 \times 10^{-4}$	$5.3 \times 10^9$	$1.89 \times 10^{-10}$	107	
T22.8	TNF- $\alpha$	$1.7 \times 10^6$	$7.5 \times 10^{-4}$	$2.3 \times 10^9$	$4.3 \times 10^{-10}$	107	
T22.9	TNF- $\alpha$	$2.3 \times 10^6$	$4.9 \times 10^{-4}$	$4.8 \times 10^9$	$2.11 \times 10^{-10}$	107	
T22.11	TNF- $\alpha$	$2.9 \times 10^6$	$7.9 \times 10^{-4}$	N/A	$2.76 \times 10^{-10}$	107	

**Example 10: Gene usage and somatic hypermutation in monoclonal antibodies**

The sequences of the heavy and kappa light chain transcripts from the described IL-8 and EGFR-human Mabs were determined Figure 6 and Figures [[ ]]. The four IL-8-specific antibodies consisted of at least three different  $V_H$  genes ( $V_{H4-34}/V_{H4-21}$ ,  $V_{H3-30}$  and  $V_{H5-51}$ ), four different  $D_H$  segments (A1/A4, K1, ir3rc, and 21-10rc) and two  $J_H$  ( $J_{H3}$  and  $J_{H4}$ ) gene segments. Three different  $V_K$  genes (012, 018, and B3) combined with  $J_K3$  and  $J_K4$  genes. Such diverse utilization shows that

Xenomouse II is capable of producing a panel of anti-IL-8 neutralizing antibodies with diverse variable regions.

In contrast to the IL-8 antibody transcripts, the sequences of antibodies selected for their ability to compete with Mab 225 showed relatively restricted  $V_H$  and  $V_K$  gene usage, with three antibodies, E1.1, E2.4 and E2.5 sharing the same  $V_H$  gene (4-31) and E2.11 containing  $V_{H4-61}$ , which is highly homologous to  $V_{H4-31}$ . Different D (2, A1/A4, XP1) and  $J_H$  ( $J_{H3}$ ,  $J_{H4}$ ,  $J_{H5}$ ) segments were detected. All four antibodies were shown to share the same  $V_K$  (018) gene. Three of them contained  $J_K4$ , and one, E2.5, contained  $J_K2$ .

Most  $V_H$  and  $V_K$  hybridoma transcripts showed extensive nucleotide changes (7-17) from the corresponding germline segments, whereas no mutations were detected in the constant regions. Most of the mutations in V segments resulted in amino acid substitutions in the predicted antibody amino acid sequences (0-12 per V gene), many in CDR1 and CDR2 regions (Figure \_). Of note are the mutations which are shared by the heavy chain sequences of EGFR antibodies, such as the Gly→Asp substitution in CDR1, shared by all antibodies, or Ser→Asn substitution in CDR2 and Val→Leu in the framework region 3 shared by three antibodies. These results indicated that an extensive process of somatic hypermutation, leading to antibody maturation and selection, is occurring in Xenomouse II.

### **Discussion**

This present application describes the first functional substitution of complex, megabase-sized mouse loci, with human DNA fragments equivalent in size and content reconstructed on YACs. With this approach, the mouse humoral immune system was "humanized" with megabase-sized human Ig loci to substantially reproduce the human antibody response in mice deficient in endogenous antibody production.

Our success in faithful reconstruction of a large portion of the human heavy and kappa light chain loci, nearly in germline configuration, establishes YAC recombination in yeast as a powerful technology to reconstitute large, complex and unstable fragments, such as the Ig loci (Mendez et al., 1995), and manipulate them for introduction into mammalian cells. Furthermore, the successful introduction of the two large heavy and kappa light chain segments into the mouse germline in intact form confirms the methodology of ES cell-yeast spheroplast fusion as a reliable and efficient approach

to delivering xenogeneic loci into the mouse germline.

Characterization of Xenomouse II strains has shown that the large Ig loci were capable of restoring the antibody system, comparable in its diversity and functionality to that of wildtype mice, and much superior to the humoral response produced in mice bearing human Ig minigene constructs (Lonberg et al., 1994) or small human Ig YACs (Green et al., 1994). This difference was manifested in the levels of mature B-cells, human Ig production, class switching efficiency, diversity, preponderance of human Ig $\kappa$  over murine Ig $\lambda$  production, and magnitude of the human antibody response, and success in the generation of high affinity, antigen-specific monoclonal antibodies to multiple antigens.

The levels of mature B-cells and human antibodies in Xenomouse II are the highest yet reported for Ig-transgenic mice, representing a several-fold increase over the levels shown for previous mice and approaching those of wildtype mice. In particular, the levels of the human IgG were more than 100 fold higher than those reported for mice bearing minilocus Ig transgenes with human  $\gamma 1$  gene (Lonberg et al., 1994). The more efficient class switching in Xenomouse II was likely the result of the inclusion of the entire switch regions, with all of their regulatory elements, as well as the additional control elements on yH2, which may be important to support and maintain proper class switching. The elevated levels of mature B-cells in Xenomouse II strains are likely to result from the higher rearrangement frequency and thus improved B-cell development in the bone marrow due to the increased V gene repertoire. B-cell reconstitution is expected to be even more pronounced in XenoMouse II strains that are homozygous for the human heavy chain locus.

The ratio of human  $\kappa$  to mouse  $\lambda$  light chain expression by circulating B-cells provides a useful internal measure of the utilization of the transgenic kappa chain locus. Whereas in mice containing one allele of smaller Ig YACs, an approximately equal distribution of human  $\kappa$  and mouse  $\lambda$  was observed, a significant preponderance of human  $\kappa$  was detected in Xenomouse II strains. Moreover, in animals homozygous for yK2 possessed a  $\kappa$ : $\lambda$  ratio that is identical to wild type mice. These observations together with the broad V $\kappa$  gene usage strongly suggest that the human proximal V $\kappa$  genes in the Xenomouse II are sufficient to support a diverse light chain response and are consistent with the bias toward proximal V $\kappa$  gene usage in humans (Cox et al., 1994).

Xenomouse II strains exhibited highly increased antibody diversity with V, D, and J genes

across the entire span of the loci accessed by the recombination mechanism and incorporated into mature antibodies. Once triggered by antigen binding, extensive somatic hypermutation occurs, leading to affinity maturation of the antibodies.

The utilization pattern of V, D, J genes in Xenomouse II also indicated they are available and utilized in a manner reminiscent of their utilization in humans, yielding an adult-like human antibody repertoire, which is different from the fetal-like, position-biased usage observed in Ig minigene-bearing mice (Taylor et al., 1992; Taylor et al., 1994; Tuailon et al., 1993). The broad utilization of many of the functional  $V_H$  and  $V_K$  genes together with the multiplicity of antigens recognized by the mice underscores the importance of the large V gene repertoire to successfully reconstituting a functional antibody response.

The ultimate test for the extent of reconstitution of the human immune response in mice is the spectrum of antigens to which the mice will elicit an antibody response and the ease with which antigen-specific high affinity Mabs can be generated to different antigens. Unlike mice engineered with smaller human Ig YACs or minigenes, which yielded to date only a limited number of antigen-specific human Mabs (Lonberg et al., 1994; Green et al., 1994; Fishwild et al., 1996), Xenomouse II generated Mabs to all human antigens tested to date. Xenomouse II strains mounted a strong human antibody response to different human antigens, presented either as soluble proteins or expressed on the surfaces of cells. Immunization with each of the three human antigens tested yielded a panel of 10-25 antigen-specific human IgG<sub>2</sub>κ Mabs. For each antigen, a set of antibodies with affinities in the range of  $10^9$ - $10^{10}$  M<sup>-1</sup> was obtained. Several measures were taken to confirm that the affinity values represent univalent binding kinetics rather than avidity: BIAcore assays with intact antibodies were carried out with sensor chips coated at low antigen density to minimize the probability of bivalent binding; for two antibodies, the assay was repeated with monovalent Fab fragments; some of the antibodies were also tested by solution radioimmunoassay. From the results of these measurements, we conclude that antibodies with affinities in the range of  $10^{10}$  M<sup>-1</sup> are readily attainable with the XenoMouse. The affinity values obtained for XenoMouse-derived antibodies are the highest to be reported for human antibodies against human antigens produced from either engineered mice (Lonberg et al., Fishwild et al., 1996) or from combinatorial libraries (Vaughan et al., 1996). These high affinities combined with the extensive amino acid substitution as a result of

somatic mutation in the V genes confirms that the mechanism of affinity maturation is intact in Xenomouse II and comparable to that in wildtype mice.

These results show that the large antibody repertoire on the human Ig YACs is being properly exploited by the mouse machinery for antibody diversification and selection, and, due to the lack of immunological tolerance to human proteins, can yield high affinity antibodies against any antigen of interest, including human antigens. The facility with which antibodies to human antigens can be generated by the human immunoglobulin genes in these mice provides further confirmation that self tolerance at the B-cell level is acquired and not inherited.

The ability to generate high affinity fully human antibodies to human antigens has obvious practical implications. Fully human antibodies are expected to minimize the immunogenic and allergic responses intrinsic to mouse or mouse-derivatized Mabs and thus to increase the efficacy and safety of the administered antibodies. Xenomouse II offers the opportunity of providing a substantial advantage in the treatment of chronic and recurring human diseases, such as inflammation, autoimmunity, and cancer, which require repeated antibody administrations. The rapidity and reproducibility with which XenoMouse II yields a panel of fully human high affinity antibodies indicates the potential advance it offers over other technologies for human antibody production. For example, in contrast to phage display, which requires intensive efforts to enhance the affinity of many of its derived antibodies and yields single chain Fvs or Fabs, Xenomouse II antibodies are high affinity fully intact immunoglobulins which can be produced from hybridomas without further engineering.

The strategy described here for creation of an authentic human humoral immune system in mice can be applied towards humanization of other multi-gene loci, such as the T cell receptor or the major histocompatibility complex, that govern other compartments of the mouse immune system (Jakobovits, 1994). Such mice would be valuable for elucidating the structure-function relationships of the human loci and their involvement in the evolution of the immune system.

#### INCORPORATION BY REFERENCE

All references cited herein, including patents, patent applications, papers, text books, and the like, and the references cited therein, to the extent that they are not already, are hereby incorporated herein by reference in their entirety. In addition, the following references are also incorporated by

reference herein in their entirety, including the references cited in such references:

Abertsen et al., "Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents." *Proc. Natl. Acad. Sci.* **87**:4256 (1990).

5

Anand et al., "Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis." *Nucl. Acids Res.* **17**:3425-3433 (1989).

Berman et al. "Content and organization of the human Ig VH locus: definition of three new V<sub>H</sub> families and linkage to the Ig C<sub>H</sub> locus." *EMBO J.* **7**:727-738 (1988).

10

Brezinschek et al., "Analysis of the heavy chain repertoire of human peripheral B-cells using single-cell polymerase chain reaction." *J. Immunol.* **155**:190-202 (1995).

Brownstein et al., "Isolation of single-copy human genes from a library of yeast artificial chromosome clones." *Science* **244**:1348- 1351 (1989).

15

Bruggemann et al. *PNAS USA* **86**:6709-6713 (1989).

Bruggemann et al., "Human antibody production in transgenic mice: expression from 100 kb of the human IgH locus." *Eur. J. Immunol.* **21**:1323-1326 (1991).

20

Bruggeman, M. and Neuberger, M.S. in *Methods: A companion to Methods in Enzymology* **2**:159-165 (Lerner et al. eds. Academic Press (1991)).

25

Bruggemann, M. and Neuberger, M.S. "Strategies for expressing human antibody repertoires in transgenic mice." *Immunology Today* **17**:391-397 (1996).

Chen et al. "Immunoglobulin gene rearrangement in B-cell deficient mice generated by targeted

deletion of the J<sub>H</sub> locus" *International Immunology* 5:647-656 (1993)

Choi et al. "Transgenic mice containing a human heavy chain immunoglobulin gene fragment cloned in a yeast artificial chromosome" *Nature Genetics* 4:117-123 (1993)

5

Coligan et al., Unit 2.1, "Enzyme-linked immunosorbent assays," in *Current protocols in immunology* (1994).

Cook, G.P. and Tomlinson, I.M., "The human immunoglobulin V<sub>H</sub> repertoire." *Immunology Today* 16:237-242 (1995).

10

Cox et al., "A directory of human germ-line Vx segments reveals a strong bias in their usage." *Eur. J. Immunol.* 24:827-836 (1994).

15 Dariavach et al., "The mouse IgH 3'-enhancer." *Eur. J. Immunol.* 21:1499-1504 (1991).

Den Dunnen et al., "Reconstruction of the 2.4 Mb human DMD-gene by homologous YAC recombination." *Human Molecular Genetics* 1:19-28 (1992).

20 Feeney, A.J. "Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences." *J. Exp. Med.* 172:137-1390 (1990).

Fishwild et al., "High-avidity human IgGκ monoclonal antibodies from a novel strain of minilocus transgenic mice." *Nature Biotech.* 14:845-851 (1996).

25

Flanagan, J.G. and Rabbitts, T.H., "Arrangement of human immunoglobulin heavy chain constant region genes implies evolutionary duplication of a segment containing γ, ε, and α genes." *Nature* 300:709-713 (1982).

Galfre, G. and Milstein, C., "Preparation of monoclonal antibodies: strategies and procedures." *Methods Enzymol.* 73:3-46 (1981).

5 Gemmill et al., "Protocols for pulsed field gel electrophoresis: Separation and detection of large DNA molecules." *Advances in Genome Biology* 1:217-251 (1991).

Gill et al., "Monoclonal anti-epidermal growth factor receptor antibodies which are inhibitors of epidermal growth factor binding and antagonists of epidermal growth factorstimulated tyrosine protein kinase activity." *J. Biol. Chem.* 259:7755 (1984).

10

Green et al., "Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs." *Nature Genetics* 7:13-21 (1994).

15 Hermanson et al., "Rescue of end fragments of yeast artificial chromosomes by homologous recombination in yeast." *Nucleic Acids Res.* 19:4943-4948 (1991).

Huber et al., "The human immunoglobulin  $\kappa$  locus. Characterization of the partially duplicated L regions." *Eur. J. Immunol.* 23:2860-2967 (1993).

20 Jakobovits, A., "Humanizing the mouse genome." *Current Biology* 4:761-763 (1994).

Jakobovits, A., "Production of fully human antibodies by transgenic mice." *Current Opinion in Biotechnology* 6:561-566 (1995).

25 Jakobovits et al., "Germ-line transmission and expression of a human-derived yeast artificial-chromosome." *Nature* 362:255-258 (1993).

Jakobovits, A. et al., "Analysis of homozygous mutant chimeric mice: Deletion of the immunoglobulin heavy-chain joining region blocks B-cell development and antibody production." *Proc. Natl. Acad.*



*Sci. USA* **90**:2551-2555 (1993).

- Kawamoto et al., "Growth stimulation of A431 cells by epidermal growth factor: Identification of high affinity receptors for EGF by an anti-receptor monoclonal antibody." *Proc. Nat. Acad. Sci., USA*  
5 **80**:1337-1341 (1983).

Lonberg et al., "Antigen-specific human antibodies from mice comprising four distinct genetic modifications." *Nature* **368**:856-859 (1994).

- 10 Lusti-Marasimhan et al., "Mutation of Leu25 and Val27 introduces CC chemokine activity into interleukin-8." *J. Biol. Chem.* **270**:2716-2721 (1995).

- Marks et al., "Oligonucleotide primers for polymerase chain reaction amplification of human immunoglobulin variable genes and design of family-specific oligonucleotide probes." *Eur. J.*  
15 *Immunol.* **21**:985-991 (1991).

Matsuda et al., "Structure and physical map of 64 variable segments in the 3' 0.8-megabase region of the human immunoglobulin heavy-chain locus." *Nature Genetics* **3**:88-94 (1993).

- 20 Max, E. *Molecular genetics of immunoglobulins. Fundamental Immunology.* 315-382 (Paul, WE, ed., New York: Raven Press (1993)).

Mendez et al., "A set of YAC targeting vectors for the interconversion of centric and acentric arms." *Cold Spring Harbor Laboratory Press, Genome Mapping and Sequencing meeting*, 163 (1993).

25

Mendez et al., "Analysis of the structural integrity of YACs comprising human immunoglobulin genes in yeast and in embryonic stem cells." *Genomics* **26**:294-307 (1995).

Ray, S. and Diamond, B., "Generation of a fusion partner to sample the repertoire of Splenic B-cells

destined for apoptosis." *Proc. Natl. Acad. Sci. USA* **91**:5548-5551 (1994).

Sato et al., "Biological effects in vitro of monoclonal antibodies to human epidermal growth factor receptors" *Mol. Biol. Med.* **1**:511-529 (1983).

5

Schiestl, R.H. and Gietz, R.D., "High efficiency transformation of intact yeast cells using stranded nucleic acids as a carrier." *Curr. Genet.* **16**:339-346 (1989).

10

Sherman et al., "Laboratory Course Manual for Methods in Yeast Genetics." (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1986)).

Silverman et al., "Meiotic recombination between yeast artificial chromosomes yields a single clone containing the entire BCL2 protooncogene." *Proc. Natl. Acad. Sci. USA* **87**:9913-9917 (19\_\_).

15

Srivastava, A. and Schlessinger, D., "Vectors for inserting selectable markers in vector arms and human DNA inserts of yeast artificial chromosomes (YACs)." *Gene* **103**:53-59 (1991).

Taylor et al., "A transgenic mouse that expresses a diversity of human sequence heavy and light chain immunoglobulins." *Nucleic Acids Research* **20**:6287-6295 (1992).

20

Taylor et al., "Human immunoglobulin transgenes undergo rearrangement, somatic mutation and class switching in mice that lack endogenous IgM." *International Immunology* **6**:579-591 (1994).

25

Tuaillon et al., "Human immunoglobulin heavy-chain minilocus recombination in transgenic mice: gene-segment use in  $\mu$  and  $\gamma$  transcripts." *Proc. Natl. Acad. Sci. USA* **90**:3720-3724 (1993).

Tuaillon et al. "Analysis of direct and inverted DJ<sub>H</sub> rearrangements in a human Ig heavy chain transgenic minilocus" *J. Immunol.* **154**:6453-6465 (1995)

Vaughan et al., "Human antibodies with subnanomolar affinities isolated from a large non-immunized phage display library." *Nature Biotech.* 14:309-314 (1996).

5 Wagner et al., "The diversity of antigen-specific monoclonal antibodies from transgenic mice bearing human immunoglobulin gene miniloci." *Eur. J. Immunol.* 24:2672-2681 (1994).

Weichhold et al., "The human immunoglobulin  $\kappa$  locus consists of two copies that are organized in opposite polarity." *Genomics* 16:503-511 (1993).

10 Yamada, M. et al., "Preferential utilization of specific immunoglobulin heavy chain diversity and joining segments in adult human peripheral blood B lymphocytes." *J. Exp. Med.* 173:395-407 (1991).

## SEQUENCE LISTING

## 1) GENERAL INFORMATION

(i) APPLICANT: Abgenix, Inc.

(ii) TITLE OF THE INVENTION: TRANSGENIC MAMMALS HAVING HUMAN IG LOCI INCLUDING PLURAL VH AND VK ...

(iii) NUMBER OF SEQUENCES: 80

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Neave

(B) STREET: 1251 Avenue of the Americas

(C) CITY: New York

(D) STATE: NY

(E) COUNTRY: USA

(F) ZIP: 10020

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE: 03-DEC-1997

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/759,620

(B) FILING DATE: 03-DEC-1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: James, Haley F

(B) REGISTRATION NUMBER: 27,794

(C) REFERENCE/DOCKET NUMBER: Cell 4.18

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 212-596-9000

(B) TELEFAX: 212-596-9090

(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGGTGCAGC TGGAGCAGTC GG

22

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCTGAGGGAG TAGAGTCCTG AGGA

24

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTACTGTGCG AGACA

15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGAGCTACG GG

12

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GACTACTGGG GC

12

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTACTGTGCG AGAGA

15

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTTGACTAC TGGGGC

16

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TTACTGTACC ACAGA

15

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 11 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACTAACTACC C

11

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTACTACTAC TACGGT

16

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTACTGTGCG AGAGA

15

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TAGGAGTGTT

10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:



(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GTAGTACCAG CTGCTAT

17

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACTACTACTA CTACGGT

17

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTACTGTGCG AGAG

14

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAGCAGCTG 10

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTTTGACTAC TGGGGC 16

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTACTGTGCG AGAGA 15

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GATATTTTGA CTGGT

15

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTACTACTAC TACGGT

16

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TTACTGTGCG AGAGA

15

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTTGACTACT GGGGC

15

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTACTGTGCG AGAGA

15

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TACTACTACT ACTACGGT

18

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATTACTGTGC GA

12

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TATAGCAGTG GCTGGT

16

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CTTTGACTAC TGGGGC

16

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE:

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TATTACTGTG CGAG

14

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GACTACTGGG GC

12

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 12 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TATTACTGTG CG

12

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GGATATAGTA GTGG

14

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs

(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTTTGACTAC TGGGGC

16

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TTACTGTGCG AGACA

15

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

ATGCTTTGAT ATCTGGGG

18

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TTAAACGAAC AGTACCCC

18

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 16 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GATCACCTTC GGCCAA

16

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE:  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ACAGGCTAAC AGTTCCCTC

20

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 14 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GGACGTTTCGG CCAA

14

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

AAGTATAACA GTGCCCC

17

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATTCATTTTC GGCCCT

16

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

ACAGTATGAT AATCTCCC

18

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GCTCACTTTC GCGGA

16

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AAAGTATAAT AGTTACCC

18

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GATCACCTTC GGCCAA

16

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CAGCATAATA GTTACCC

17

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATTCACCTTC GGCCCT

16

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

AATATTATAG TACTCC

16

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GCTCACTTTC GCGGA

16

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CAGTATGGTA GCTCACCTC

19

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 0 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

CACTTTTGGCCAG

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Ala	Ser	Thr	Lys	Gly	Pro	Ser	Val	Phe	Pro	Leu	Ala	Pro	Cys	Ser	Arg
1		5		10						15					
Ser	Thr	Ser	Thr												
															20

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Leu	Ser	Leu	Thr	Cys	Ala	Val	Tyr	Gly	Gly	Ser	Phe	Ser	Gly	Tyr	Tyr
1		5		10						15					
Trp	Ser	Trp	Ile	Arg	Gln	Pro	Pro	Gly	Lys	Gly	Leu	Glu	Trp	Ile	Gly
		20		25						30					
Glu	Ile	Asn	His	Ser	Gly	Ser	Thr	Asn	Tyr	Asn	Pro	Ser	Leu	Lys	Ser
		35		40						45					
Arg	Val	Thr	Ile	Ser	Val	Asp	Thr	Ser	Lys	Asn	Gln	Phe	Ser	Leu	Lys
		50		55						60					
Leu	Ser	Ser	Val	Thr	Ala	Ala	Asp	Thr	Ala	Val	Tyr	Tyr	Cys	Ala	Arg
65				70				75							80

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 119 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

```

Leu Ser Leu Thr Cys Ala Val Tyr Gly Gly Ser Phe Ser Gly Tyr Tyr
1      5      10      15
Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile Gly
20     25     30
Glu Ile Asn Gln Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser
35     40     45
Arg Val Ile Ile Ser Ile Asp Thr Ser Lys Thr Gln Phe Ser Leu Lys
50     55     60
Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg
65     70     75     80
Glu Thr Pro His Ala Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr
85     90     95
Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro
100    105    110
Cys Ser Arg Ser Thr Ser Thr
115

```

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

```

Leu Ser Leu Thr Cys Ala Val Tyr Gly Gly Ser Phe Ser Gly Tyr Tyr
1      5      10      15
Trp Thr Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile Gly
20     25     30
Glu Ile Ile His His Gly Asn Thr Asn Tyr Asn Pro Ser Leu Lys Ser
35     40     45

```

Arg Val Ser Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu Thr  
 50                55                60  
 Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg  
 65                70                75                80  
 Gly Gly Ala Val Ala Ala Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val  
 85                90                95  
 Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala  
 100                105                110  
 Pro Cys Ser Arg Ser Thr Ser Thr  
 115                120

## (2) INFORMATION FOR SEQ ID NO:55:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ser His His Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr  
 1                5                10                15  
 Ser Tyr Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu  
 20                25                30  
 Trp Met Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro  
 35                40                45  
 Ser Phe Gln Gly Gln Val Thr Ile Ser Ala Asp Lys Ser Ile Ser Thr  
 50                55                60  
 Ala Tyr Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr  
 65                70                75                80  
 Tyr Cys Ala Arg

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 121 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr  
 1 5 10 15  
 Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met  
 20 25 30  
 Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe  
 35 40 45  
 Gln Gly Gln Val Thr Ile Ser Ala Asp Lys Ser Ile Ser Thr Ala Tyr  
 50 55 60  
 Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys  
 65 70 75 80  
 Ala Arg Gln Asp Gly Asp Ser Phe Asp Tyr Trp Gly Gln Gly Thr Leu  
 85 90 95  
 Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu  
 100 105 110  
 Ala Pro Cys Ser Arg Ser Thr Ser Thr  
 115 120

## (2) INFORMATION FOR SEQ ID NO:57:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 83 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser  
 1 5 10 15  
 Tyr Gly Met His Trp Xaa Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp  
 20 25 30  
 Val Ala Val Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Ala Asp Ser  
 35 40 45  
 Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu  
 50 55 60  
 Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr  
 65 70 75 80  
 Cys Ala Arg

## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 122 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide



- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```

Arg Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser
 1      5      10      15
Tyr Gly Met His Trp Xaa Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp
 20     25     30
Val Ala Glu Ile Ser Tyr Asp Gly Ser Asn Lys Tyr Tyr Val Asp Ser
 35     40     45
Val Lys Gly Arg Leu Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu
 50     55     60
Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr
 65     70     75     80
Cys Ala Arg Asp Arg Leu Gly Ile Phe Asp Tyr Trp Gly Gln Gly Thr
 85     90     95
Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro
 100    105    110
Leu Ala Pro Cys Ser Arg Ser Thr Ser Thr
 115    120

```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
 1      5      10      15
Gln

```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp
 1       5       10      15
Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Asp Ala
 20      25      30
Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser
 35      40      45
Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile
 50      55      60
Ala Thr Tyr Tyr Cys Gln Gln Asp Asn Leu Pro
65      70      75

```

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Lys Phe Leu Ser Trp
 1       5       10      15
Phe Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Gly Thr
 20      25      30
Ser Tyr Leu Glu Thr Gly Val Pro Ser Ser Phe Ser Gly Ser Gly Ser
 35      40      45
Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Val
 50      55      60
Ala Thr Tyr Phe Cys Gln Gln Asp Asp Leu Pro Tyr Thr Phe Gly Pro
65      70      75      80
Gly Thr Lys Val Asp Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
 85      90      95
Ile Phe Pro Pro Ser Asp Glu Gln
100

```

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

```

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp
1       5       10      15
Tyr Gln Gln Lys Ala Gly Lys Ala Pro Lys Val Leu Ile Tyr Ala Ala
20      25      30
Ser Asn Leu Glu Ala Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser
35      40      45
Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile
50      55      60
Ala Thr Tyr Tyr Cys His Gln Asp Asn Leu Pro Leu Thr Phe Gly Gly
65      70      75      80
Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe
85      90      95
Ile Phe Pro Pro Ser Asp Glu Gln
100

```

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 74 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

```

Thr Cys Arg Ala Ser Gln Ser Ile Ser Ser Tyr Leu Asn Trp Tyr Gln
1       5       10      15
Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Ala Ala Ser Ser
20      25      30
Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr
35      40      45
Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr
50      55      60
Tyr Tyr Cys Gln Gln Ser Tyr Ser Thr Pro
65      70

```

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 86 amino acids
- (B) TYPE: amino acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE: internal  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```

Thr Cys Arg Ala Ser Gln Ser Ile Ser Asn Tyr Leu Asn Trp Tyr Gln
 1      5      10      15
Gln Lys Pro Gly Lys Ala Pro Lys Phe Leu Ile Tyr Gly Ala Ser Ser
 20     25     30
Leu Glu Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr
 35     40     45
Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr
 50     55     60
Tyr Tyr Cys Gln Gln Ser Tyr Ser Asn Pro Leu Thr Phe Gly Gly Gly
 65     70     75     80
Thr Lys Val Glu Ile Lys
      85

```

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
(iii) HYPOTHETICAL: NO  
(iv) ANTISENSE: NO  
(v) FRAGMENT TYPE: internal  
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

```

Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Leu Tyr Ser Ser Asn Asn
 1      5      10      15
Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys
 20     25     30
Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val Pro Asp Arg
 35     40     45
Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
 50     55     60
Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Ser
 65     70     75     80
Thr Pro

```

(2) INFORMATION FOR SEQ ID NO:66:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 94 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

```

Thr Ile Asn Cys Lys Ser Ser Gln Ser Val Leu Tyr Ile Ser Asn Asn
 1       5       10      15
Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys
      20      25      30
Leu Leu Ile Tyr Trp Ala Ser Thr Arg Lys Ser Gly Val Pro Asp Arg
      35      40      45
Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser
      50      55      60
Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln Tyr Tyr Asp
      65      70      75      80
Thr Pro Phe Thr Phe Gly Pro Gly Thr Lys Val Asp Ile Lys
      85      90

```

## (2) INFORMATION FOR SEQ ID NO:67:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

```

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg
 1       5       10      15
Ser Thr Ser Thr
      20

```

## (2) INFORMATION FOR SEQ ID NO:68:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```

Val Ser Gly Gly Ser Ile Ser Ser Gly Gly Tyr Tyr Trp Ser Trp Ile
 1      5      10      15
Arg Gln His Pro Gly Lys Gly Leu Glu Trp Ile Gly Tyr Ile Tyr Tyr
 20     25     30
Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile
 35     40     45
Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val
 50     55     60
Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg
 65     70     75

```

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 118 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: peptide
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE: internal
- (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```

Val Ser Gly Gly Ser Ile Asn Ser Gly Asp Tyr Tyr Trp Ser Trp Ile
 1      5      10      15
Arg Gln His Pro Gly Lys Gly Leu Asp Cys Ile Gly Tyr Ile Tyr Tyr
 20     25     30
Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile
 35     40     45
Ser Val Asp Thr Ser Lys Asn Gln Phe Phe Leu Lys Leu Thr Ser Val
 50     55     60
Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Ser Thr Val Val
 65     70     75     80
Asn Pro Gly Trp Phe Asp Pro Trp Gly Gln Gly Tyr Leu Val Thr Val
 85     90     95
Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys
100     105     110
Ser Arg Ser Thr Ser Thr
115

```

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 117 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTISENSE: NO  
 (v) FRAGMENT TYPE: internal  
 (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

```

Val Ser Gly Gly Ser Ile Asn Ser Gly Asp Tyr Tyr Trp Ser Trp Ile
 1       5       10      15
Arg Gln His Pro Gly Lys Gly Leu Glu Trp Ile Gly Ser Ile Tyr Tyr
 20      25      30
Ser Gly Asn Thr Phe Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile
 35      40      45
Ser Leu Asp Thr Ser Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val
 50      55      60
Thr Ala Ala Asp Thr Ala Val Cys Tyr Cys Ala Arg Asn Ile Val Thr
 65      70      75      80
Thr Gly Ala Phe Asp Ile Trp Gly Gln Gly Thr Met Val Thr Val Ser
 85      90      95
Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser
100     105     110
Arg Ser Thr Ser Thr
115
  
```

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 116 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
 (iii) HYPOTHETICAL: NO  
 (iv) ANTISENSE: NO  
 (v) FRAGMENT TYPE: internal  
 (vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```

Val Ser Gly Gly Ser Ile Ser Ser Gly Asp Tyr Tyr Trp Thr Trp Ile
 1       5       10      15
Arg Gln His Pro Gly Lys Gly Leu Glu Trp Ile Gly Tyr Ile Tyr Tyr
 20      25      30
Ser Gly Asn Thr Tyr Tyr Asn Pro Ser Leu Lys Ser Arg Val Ser Met
 35      40      45
Ser Ile Asp Thr Ser Glu Asn Gln Phe Ser Leu Lys Leu Ser Ser Val
 50      55      60
Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Lys Pro Val Thr
  
```

65            70            75            80  
 Gly Gly Glu Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser  
              85            90            95  
 Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser Arg  
              100            105            110  
 Ser Thr Ser Thr  
 115

## (2) INFORMATION FOR SEQ ID NO:72:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Val Ser Gly Gly Ser Val Ser Ser Gly Ser Tyr Tyr Trp Ser Trp Ile  
 1            5            10            15  
 Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile Gly Tyr Ile Tyr Tyr  
              20            25            30  
 Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile  
              35            40            45  
 Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val  
              50            55            60  
 Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg  
 65            70            75

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Val Ser Gly Gly Ser Val Ser Ser Gly Ser Tyr Tyr Trp Ser Trp Ile  
 1            5            10            15  
 Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile Gly Tyr Ile Tyr Tyr  
              20            25            30



Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile  
 35 40 45  
 Ser Val Asp Thr Ser<sup>1</sup> Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val  
 50 55 60  
 Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg  
 65 70 75

## (2) INFORMATION FOR SEQ ID NO:74:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Val Ser Gly Gly Ser Val Ser Ser Gly Asp Tyr Tyr Trp Ser Trp Ile  
 1 5 10 15  
 Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile Gly His Leu Tyr Tyr  
 20 25 30  
 Ser Gly Asn Thr Asn Tyr Asn Pro Ser Leu Lys Ser Arg Val Thr Ile  
 35 40 45  
 Ser Leu Asp Thr Ser Lys Asn Gln Phe Ser Leu Lys Leu Ser Ser Val  
 50 55 60  
 Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala Arg Asp Phe Leu Thr  
 65 70 75 80  
 Gly Ser Phe Phe Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser  
 85 90 95  
 Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Cys Ser  
 100 105 110  
 Arg Ser Thr Ser Thr  
 115

## (2) INFORMATION FOR SEQ ID NO:75:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu  
 1 5 10 15  
 Gln

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp  
 1 5 10 15  
 Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Asp Ala  
 20 25 30  
 Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser  
 35 40 45  
 Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile  
 50 55 60  
 Ala Thr Tyr Tyr Cys Gln Gln Asp Asn Leu Pro  
 65 70 75

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Asn Asn Tyr Leu Asn Trp  
 1 5 10 15  
 Phe Gln Gln Lys Pro Gly Lys Ala Pro Lys Val Leu Ile His Asp Ala  
 20 25 30  
 Ser Asn Leu Glu Thr Gly Gly Pro Ser Arg Phe Ser Gly Ser Gly Ser  
 35 40 45  
 Gly Thr Asp Phe Thr Phe Thr Ile Ser Gly Leu Gln Pro Glu Asp Ile

50            55            60  
 Ala Thr Tyr Tyr Cys Gln Gln Glu Ser Leu Pro Leu Thr Phe Gly Gly  
 65            70            75            80  
 Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe  
           85            90            95  
 Ile Phe Pro Pro Ser Asp Glu Gln  
 100

## (2) INFORMATION FOR SEQ ID NO:78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Thr Ile Tyr Leu Asn Trp  
 1            5            10            15  
 Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Asn Asp Ala  
 20            25            30  
 Ser Ser Leu Glu Thr Gly Val Pro Leu Arg Phe Ser Gly Ser Gly Ser  
 35            40            45  
 Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Ile  
 50            55            60  
 Ala Thr Tyr Tyr Cys Gln Gln Asp His Leu Pro Leu Thr Phe Gly Gly  
 65            70            75            80  
 Gly Thr Lys Val Ala Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe  
 85            90            95  
 Ile Phe Pro Pro Ser Asp Glu Gln  
 100

## (2) INFORMATION FOR SEQ ID NO:79:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp  
 1 5 10 15  
 Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Asp Ala  
 20 25 30  
 Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser  
 35 40 45  
 Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro Glu Asp Val  
 50 55 60  
 Gly Thr Tyr Tyr Val Gln Gln Glu Ser Leu Pro Cys Gly Phe Gly Gln  
 65 70 75 80  
 Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe  
 85 90 95  
 Ile Phe Pro Pro Ser Asp Glu Gln  
 100

## (2) INFORMATION FOR SEQ ID NO:80:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (iii) HYPOTHETICAL: NO

## (iv) ANTISENSE: NO

## (v) FRAGMENT TYPE: internal

## (vi) ORIGINAL SOURCE:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr Leu Asn Trp  
 1 5 10 15  
 Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Asn Asp Ala  
 20 25 30  
 Ser Asp Leu Glu Thr Gly Val Pro Ser Arg Ile Ser Gly Ser Gly Ser  
 35 40 45  
 Gly Thr Asp Phe Thr Phe Thr Ile Ser Asn Leu Gln Pro Glu Asp Ile  
 50 55 60  
 Ala Thr Tyr Tyr Cys Gln Gln Asp Ser Leu Pro Leu Thr Phe Gly Gly  
 65 70 75 80  
 Gly Thr Lys Val Glu Ile Arg Arg Thr Val Ala Ala Pro Ser Val Phe  
 85 90 95  
 Ile Phe Pro Pro Ser Asp Glu Gln  
 100

## CLAIMS

### WHAT IS CLAIMED IS:

1. A transgenic non-human mammal having a genome, the genome comprising  
5 modifications, the modifications comprising:
  - an inactivated endogenous immunoglobulin (Ig) locus, such that the mammal would not display normal B-cell development;
  - an inserted human heavy chain Ig locus in substantially germline configuration, the human heavy chain Ig locus comprising a human mu constant region and regulatory and  
10 switch sequences thereto, a plurality of human J<sub>H</sub> genes, a plurality of human D<sub>H</sub> genes, and a plurality of human V<sub>H</sub> genes; and
  - an inserted human kappa light chain Ig locus in substantially germline configuration, the human kappa light chain Ig locus comprising a human kappa constant region, a plurality of J<sub>κ</sub> genes, and a plurality of V<sub>κ</sub> genes,
  - 15 wherein the number of V<sub>H</sub> and V<sub>κ</sub> genes inserted are selected to substantially restore normal B-cell development in the mammal.
2. The mammal of Claim 1, wherein the heavy chain Ig locus comprises a second constant region selected from the group consisting of human gamma-1, human gamma-2, human gamma-3, human gamma-4, alpha, epsilon, and delta.
- 20 3. The mammal of Claim 1, wherein the number of V<sub>H</sub> genes is greater than about 20.
4. The mammal of Claim 1, wherein the number of V<sub>κ</sub> genes is greater than about 15.
5. The mammal of Claim 1, wherein the number of D<sub>H</sub> genes is greater than about 25, the number of J<sub>H</sub> genes is greater than about 4, the number of V<sub>H</sub> genes is greater than about 20, the number of J<sub>κ</sub> genes is greater than about 4, and the number of V<sub>κ</sub> genes is greater than about 15.
- 25 6. The mammal of Claim 1, wherein the number of D<sub>H</sub> genes, the number of J<sub>H</sub> genes, the number of V<sub>H</sub> genes, the number of J<sub>κ</sub> genes, and the number of V<sub>κ</sub> genes are selected such that the Ig loci are capable of encoding greater than about  $1 \times 10^5$  different functional antibody sequence combinations.
7. The mammal of Claim 1, wherein in a population of mammals B-cell function is

reconstituted on average to greater than about 50% as compared to wild type.

8. In a transgenic non-human mammal having a genome that comprises modifications, the modifications rendering the mammal capable of producing human immunoglobulin molecules but substantially incapable of producing functional endogenous antibody molecules, the improvement comprising:

insertion into the genome of the mammal of sufficient human  $V_H$ ,  $D_H$ ,  $J_H$ ,  $V_K$ , and  $J_K$  genes such that the mammal is capable encoding greater than about  $1 \times 10^6$  different functional human immunoglobulin sequence combinations, without accounting for junctional diversity or somatic mutation events.

9. In a transgenic non-human mammal having a genome that comprises modifications, the modifications rendering the mammal capable of producing human immunoglobulin molecules but substantially incapable of producing functional endogenous antibody molecules, which modifications, with respect to the mammal's incapacity to produce functional endogenous antibody molecules would not allow the mammal to display normal B-cell development, the improvement comprising:

insertion into the genome of the mammal of sufficient human  $V_H$ ,  $D_H$ ,  $J_H$ ,  $V_K$ , and  $J_K$  genes such that the mammal is capable of encoding greater than about  $1 \times 10^6$  different functional human immunoglobulin sequence combinations and sufficient  $V_H$  and  $V_K$  genes to substantially restore normal B-cell development in the mammal.

10. In the mammal of Claim 9, wherein in a population of mammals B-cell function is reconstituted on average to greater than about 50% as compared to wild type.

11. A transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising:

an inactivated endogenous heavy chain immunoglobulin (Ig) locus;

an inactivated endogenous kappa light chain Ig locus;

an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2; and

an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

12. A transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising:

an inactivated endogenous heavy chain immunoglobulin (Ig) locus;

an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2; and

an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

13. A transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising:

an inactivated endogenous heavy chain immunoglobulin (Ig) locus;

an inactivated endogenous kappa light chain Ig locus;

an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2 without the presence of a human gamma-2 constant region; and

an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

14. A transgenic non-human mammal having a genome, the genome comprising modifications, the modifications comprising:

an inactivated endogenous heavy chain immunoglobulin (Ig) locus;

an inactivated endogenous kappa light chain Ig locus;

an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yH2 without the presence of a human gamma-2 constant region; and

an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of yK2.

15. A transgenic non-human mammal having a genome, the genome comprising

modifications, the modifications comprising:

an inactivated endogenous heavy chain immunoglobulin (Ig) locus;

an inserted human heavy chain Ig locus, the human heavy chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of  $\gamma H2$  without the presence of a human gamma-2 constant region; and

an inserted human kappa light chain Ig locus, the human kappa light chain Ig locus comprising a nucleotide sequence substantially corresponding to the nucleotide sequence of  $\kappa K2$ .

16. A method for the production of human antibodies, comprising:

inoculating a mammal of any one of Claims 1-10 with an antigen;

collecting and immortalizing lymphocytic cells to obtain immortal cell lines secreting human antibodies that specifically bind to the antigen with an affinity of greater than  $10^9 M^{-1}$ ; and

isolating the antibodies from the immortal cell lines.

17. The method of Claim 11, wherein the antigen is IL-8.

18. The method of Claim 11, wherein the antigen is EGFR.

19. The method of Claim 11, wherein the antigen is TNF- $\alpha$ .

20. An antibody produced by the method of Claim 11.

21. An anti-IL-8 antibody produced by the method of Claim 12.

22. An anti-EGFR antibody produced by the method of Claim 13.

23. An anti-TNF- $\alpha$  antibody produced by the method of Claim 14.

24. In a method for the production of transgenic mice, the transgenic mice having a genome, the genome comprising modifications, the modifications comprising insertion of a plurality of human variable regions, the improvement comprising:

insertion of the human variable regions from a yeast artificial chromosome.

25. Transgenic mice and transgenic offspring therefrom produced through use of the improvement of Claim 19.

26. In a transgenic mammal, the transgenic mammal comprising a genome, the genome comprising modifications, the modifications comprising an inserted human heavy chain



immunoglobulin transgene, the improvement comprising:

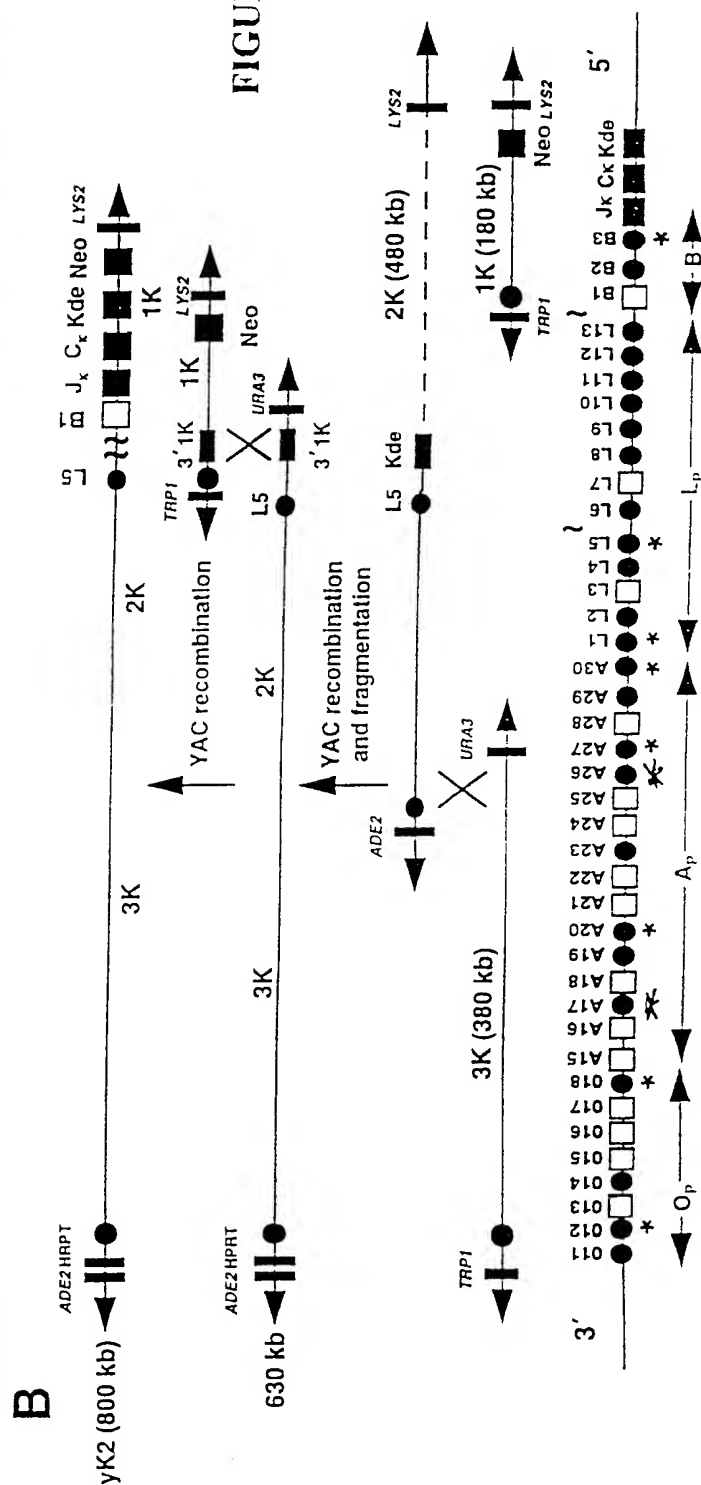
the transgene comprising selected sets of human variable region genes that enable human-like junctional diversity and human-like complementarity determining region 3 (CDR3) lengths.

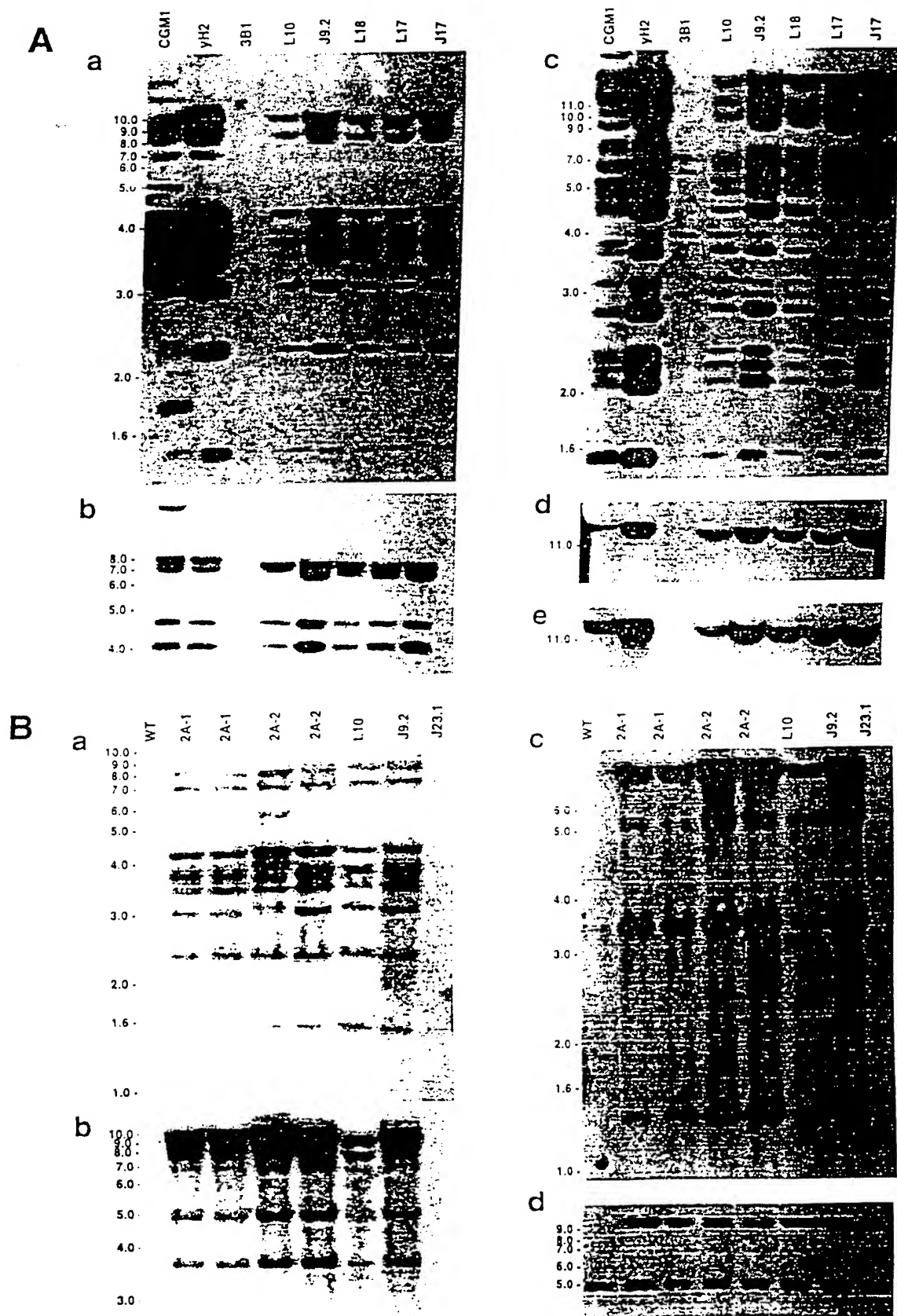
5        27.     In the improvement of Claim 26, wherein the human-like junctional diversity comprises average N-addition lengths of 7.7 bases.

      28.     In the improvement of Claim 26, wherein the human-like CDR3 lengths comprise between about 2 through about 25 residues with an average of about 14.



# FIGURE 1





**FIGURE 2**

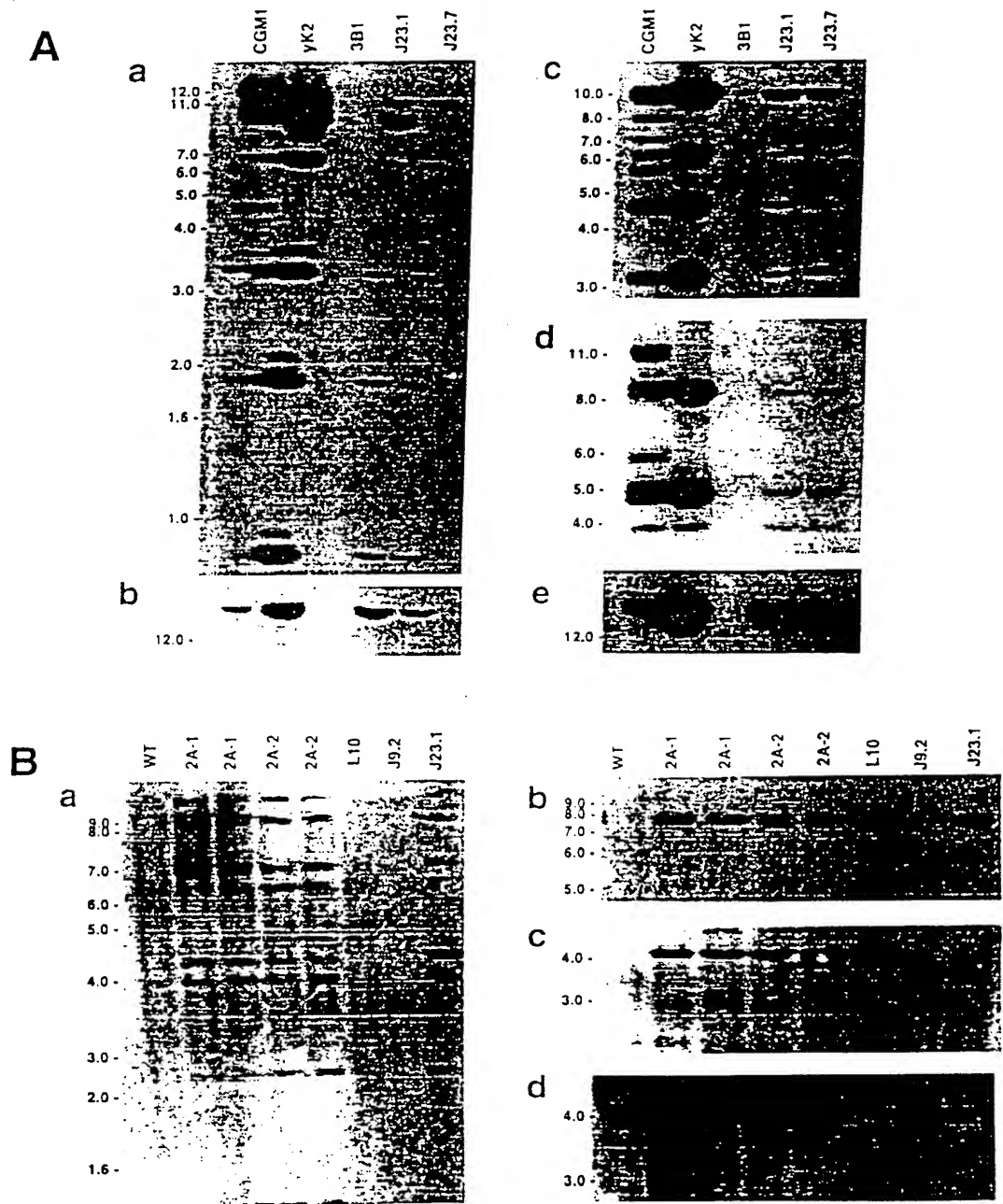


FIGURE 3

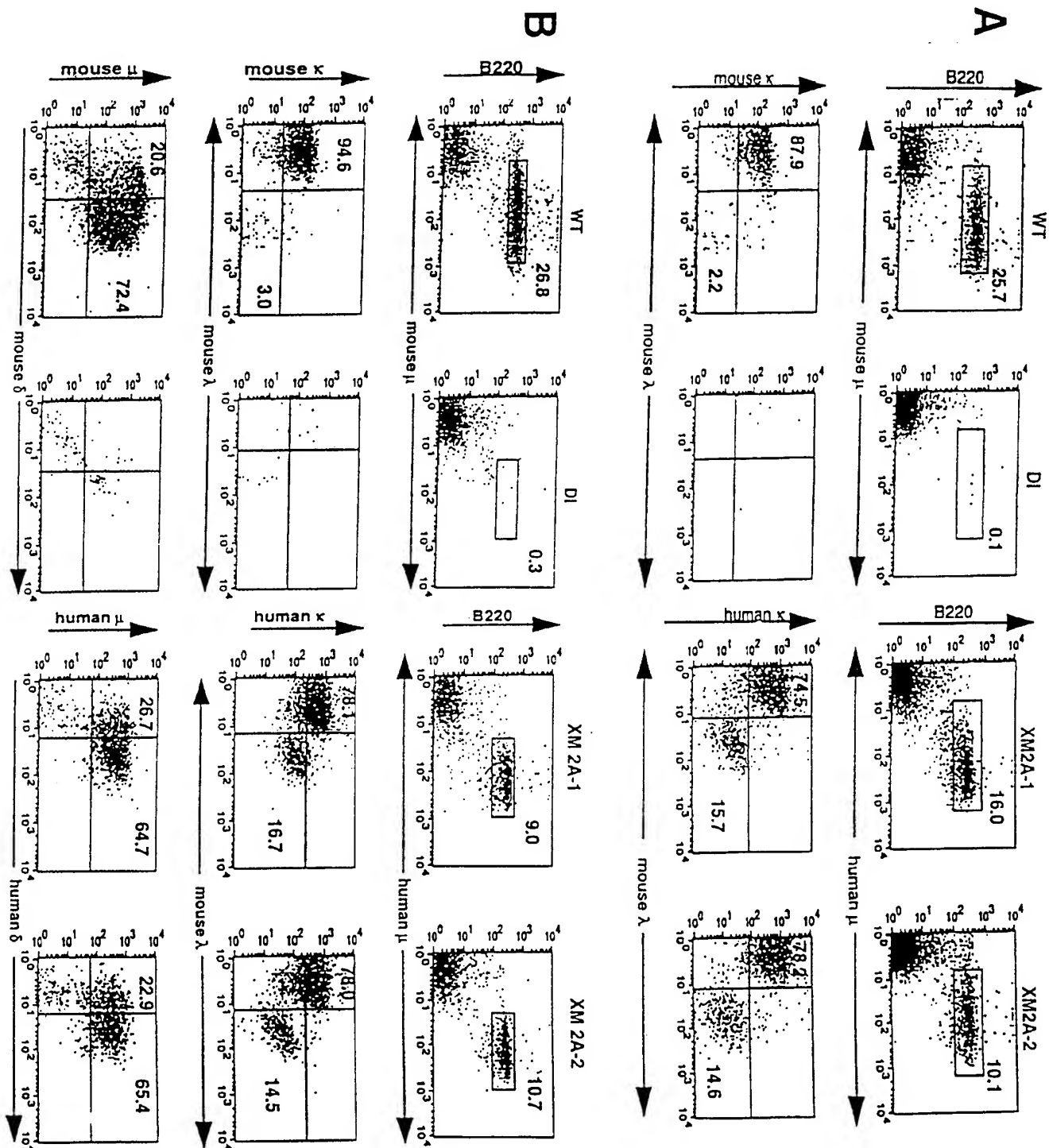
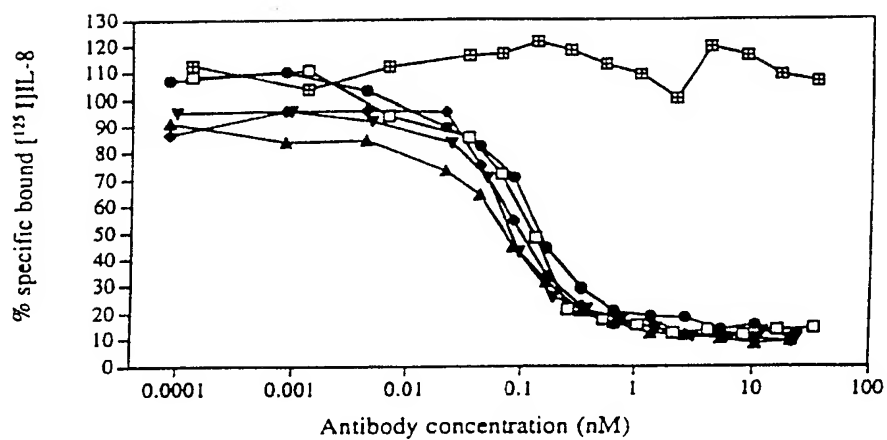
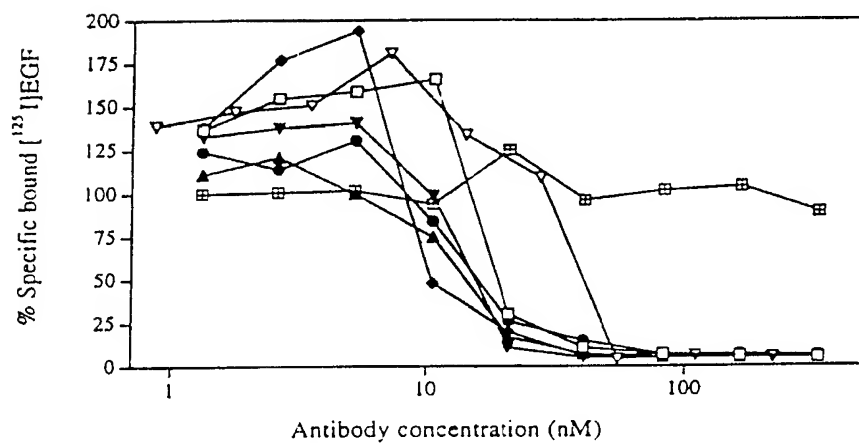
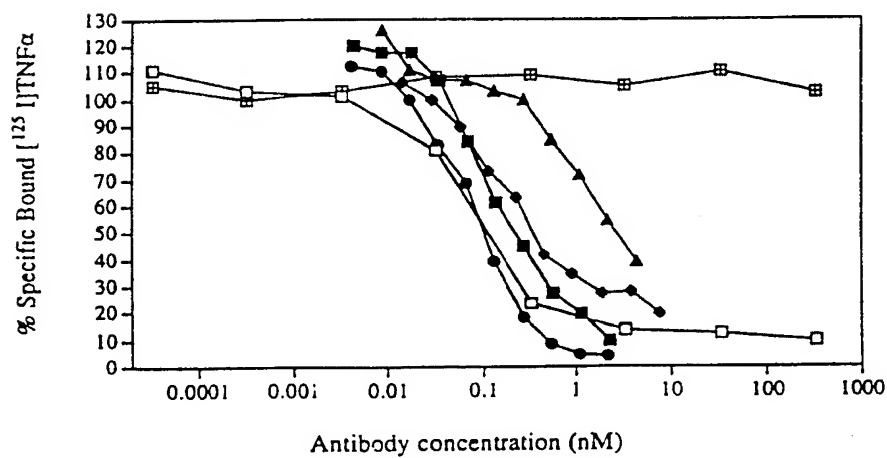


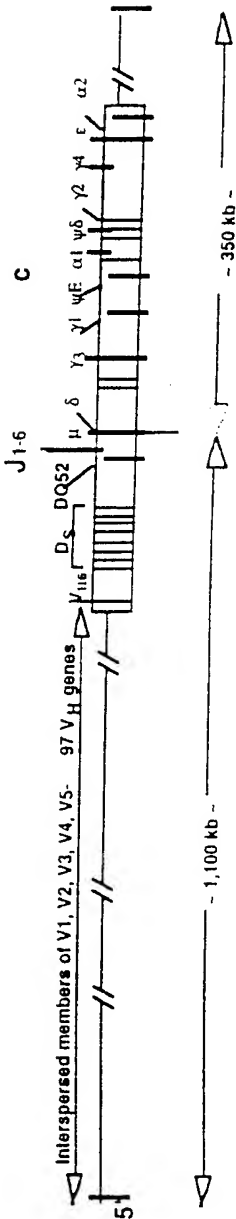
FIGURE 4

**A****B****C****FIGURE 5**



# Human Ig Genes Reside on Large, Complex Loci

Human Heavy Chain Locus - Chromosome 14



Human Kappa Chain Locus - Chromosome 2

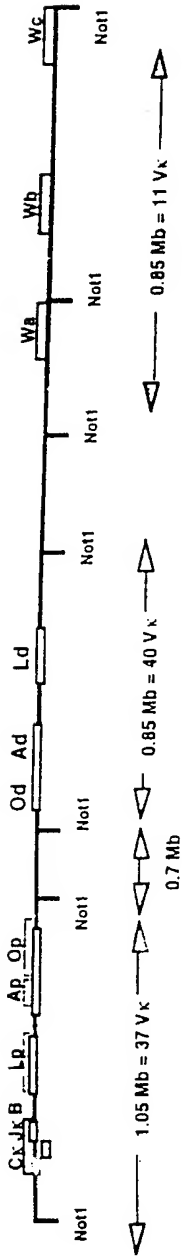


FIGURE 7





# Reconstruction of Human Kappa Chain Proximal Locus on a Single YAC

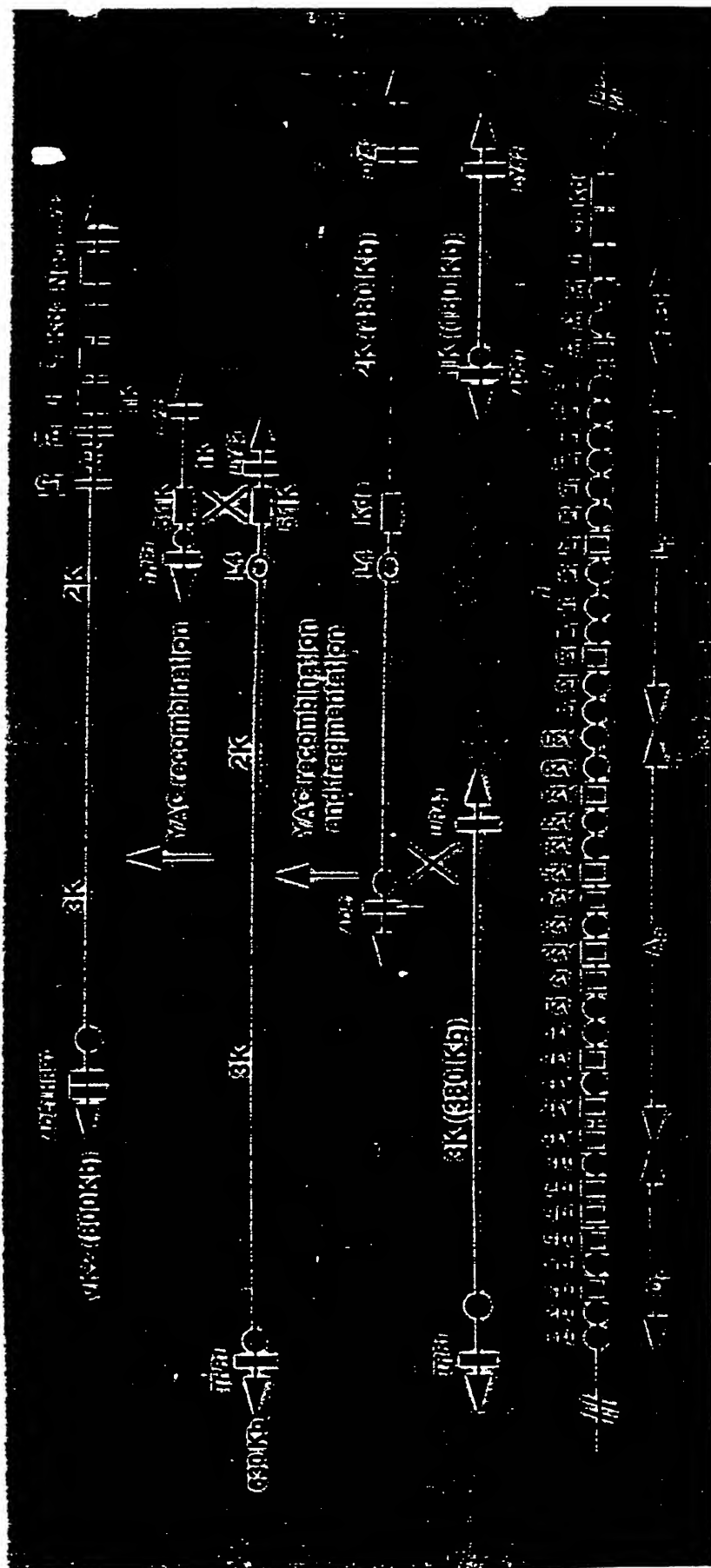
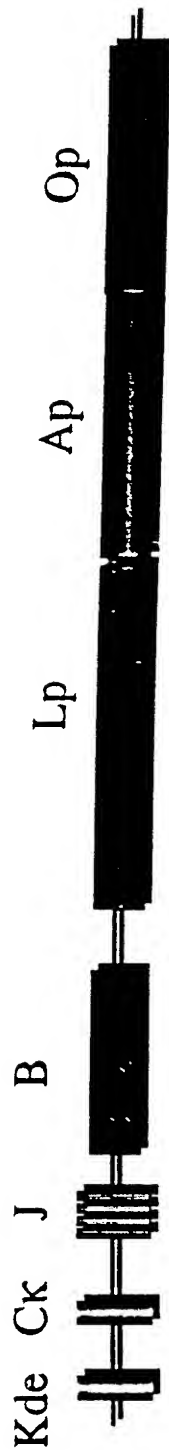


FIGURE 9

# Phase II Xenomouse: Human Kappa Chain



- 800 kb YAC
- ~32 Vκ genes

FIGURE 10

# 1020 kb Heavy Chain YAC is Integrated Intact into the Mouse Genome

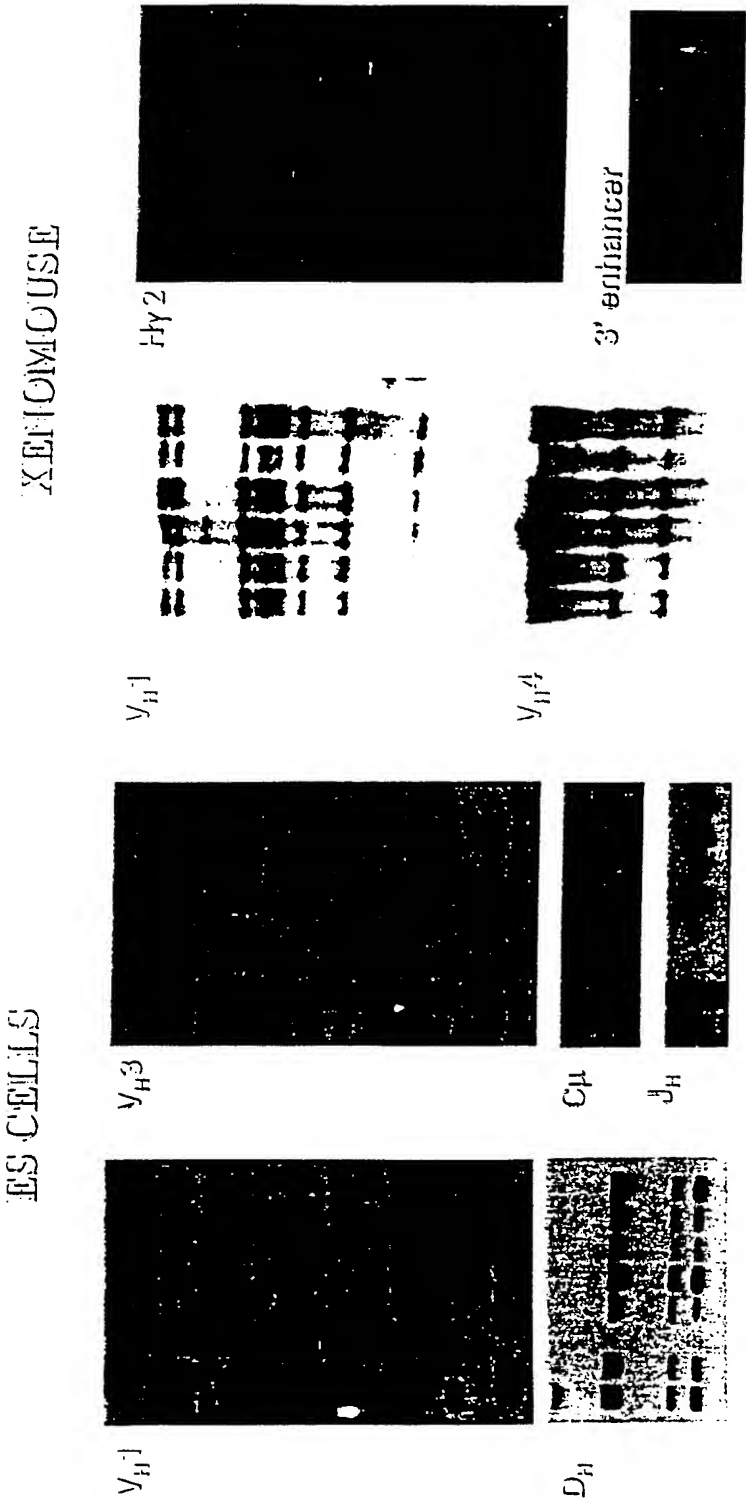


FIGURE 11

# 800 kb Kappa Light Chain YAC is Integrated Intact into the Mouse Genome

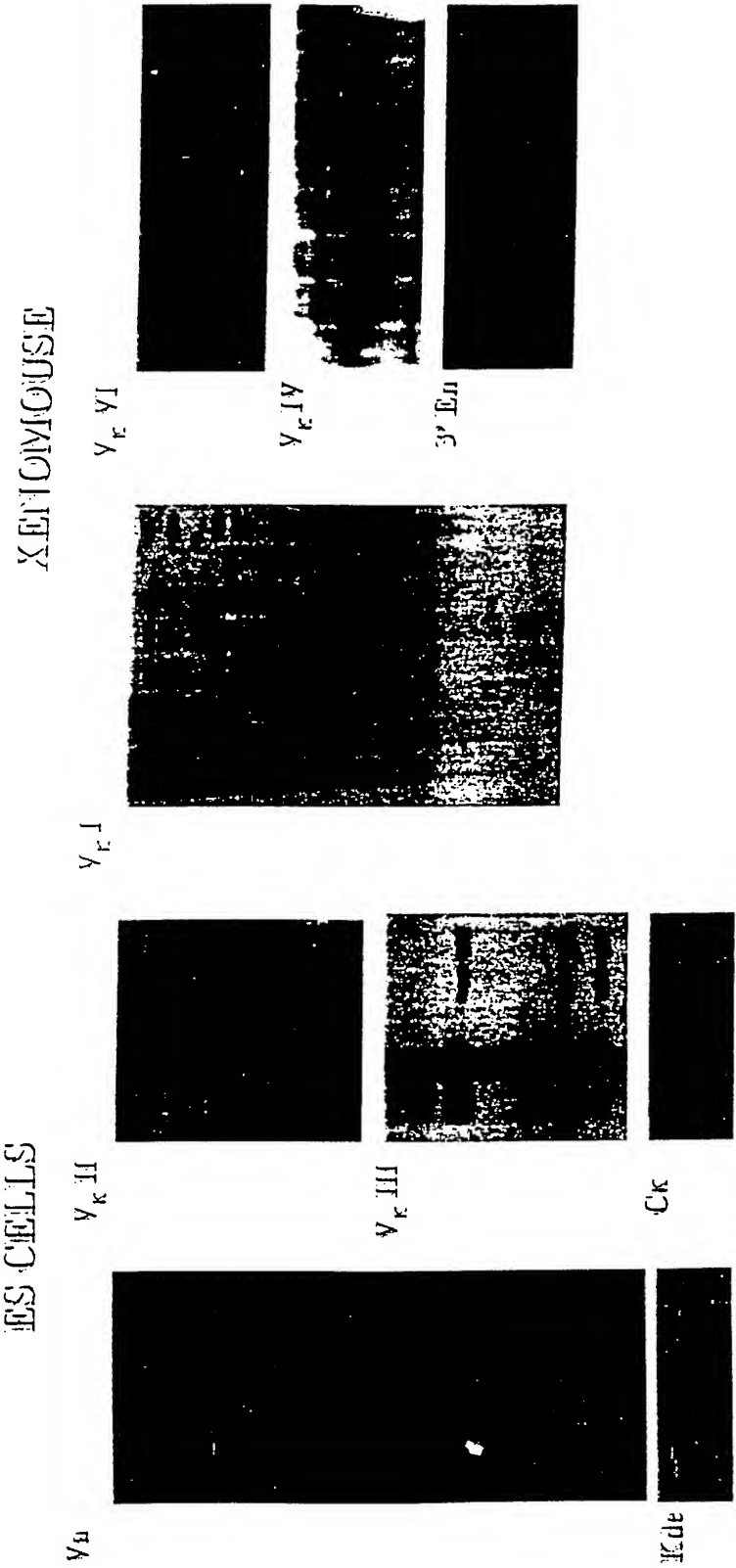


FIGURE 12

# B Cell Reconstitution In XenoMouse Blood

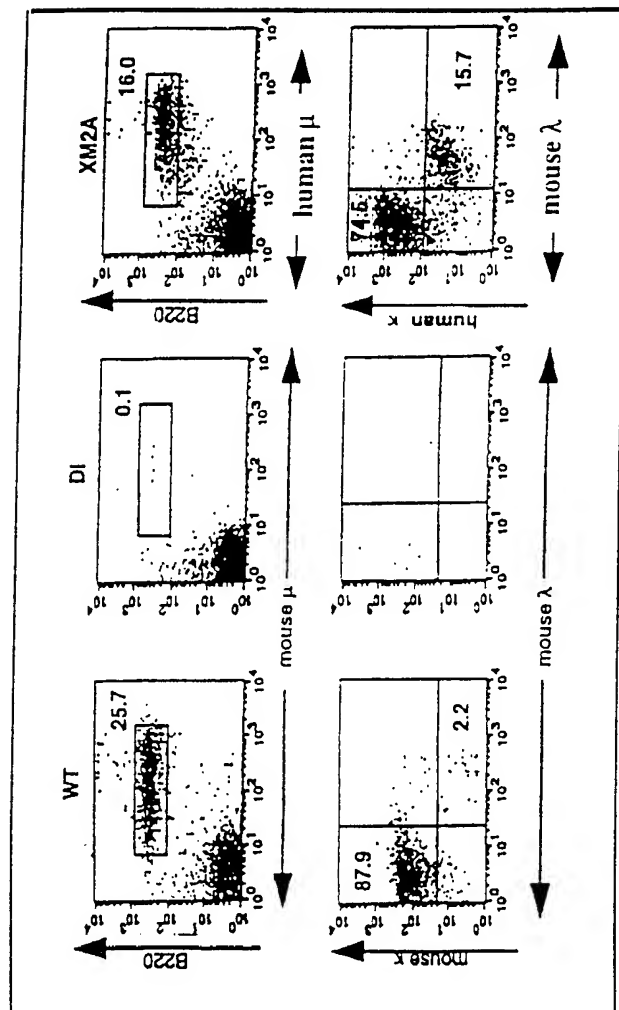


FIGURE 13

# Production of Human Antibodies by XenoMouse II

		XenoMouse*	Wild Type Mouse (B6x129)*
		<u>(Human Ig <math>\mu</math>g/<math>\mu</math>l)</u>	<u>(Mouse Ig <math>\mu</math>g/<math>\mu</math>l)</u>
$\mu$	700	400	
$\gamma$	600 (h $\gamma$ 2)	2000 (all $\gamma$ isotypes)	
$\kappa$	800	2000	

\* kept in pathogen-free conditions

FIGURE 14

# Repertoire Analysis of Human Heavy Chain Transcripts Expressed in XenoMouse II Strains

Clone		V <sub>H</sub>	N	D <sub>H</sub>	N	J <sub>H</sub>
A2.1	5-51 (DP73)	TTACTGTGCGAGACA	4 (TAGG)	XP5ic	AATCAT	JH4
B2.1.5	3-33 (DP-50)	TTACTGTGCGAGAGA	7 (TCGGGGA)	3ic	AATAGCA	JH4
B4.2.4	3-15 (DP-38)	TTACTGTACCACAGA	1 (G)	K1	GGCTAC	JH6
B4.2.5	4-59 (DP-71)	TTACTGTGCGAGAGA	10 (TAGGAGTGT)	4	GIAGTACCAGCTGCTAT	JH6
D2.2.5	4-34 (DP-63)	TTACTGTGCGAGAG_	2 (GG)	N1	GCAGCAGCTG	JH4
D2.1.3	3-48 (DP51)	TTACTGTGCGAGAGA	4 (TCTT)	XP1	GATATTTTGACTGGT	JH6
D2.2.8	4-31 (DP-65)	TTACTGTGCGAGAGA	2 (GA)	A4	GACTGCAG	JH4
A2.2.4	3-21 (DP-77)	TTACTGTGCGAGAGA	2 (TT)	IR3	GGGGCTGG	JH6
D4.2.11	4-4/35	ATTACTGTGCGA	1 (A)	N1	TATAGCAGTGGCTGGT	JH4
C1.2.1	1-18 (DP-14)	TATTACTGTGCGAG_	0	XP1/21-7	GTTA	JH4
C3.1.2	4-39 (DP-79)	TATTACTGTGCG_	3 (GCC)	2	GGATATAGTAGTGG	JH4
D2.2.7	5-51 (DP73)	TTACTGTGCGAGACA	4 (TGGC)	K1	AGTGGCT	JH3

FIGURE 15



# Repertoire Analysis of Human Kappa Light Chain Transcripts Expressed in XenoMouse II Strains

Clone	V <sub>K</sub>	N	J <sub>K</sub>
F2.2.3	02(DPK9)	0	JK5
F4.1.8	L5/(DPK5)	0	JK1
F4.1.6	A20/(DPK4)	0	JK3
F2.2.5	08	0	JK4
F2.1.5	L1	0	JK5
F2.1.4	A30	0	JK3
F2.1.3	B3/(DPK24)	0	JK4
F4.1.3	A27(DPK22)	1(G)	JK2

FIGURE 16



# High Antigen-Specific Human Antibody Response in XenoMouse

<u>Antigen</u>	<u>Titer</u>		
	<u>hIgM</u>	<u>hIgG</u>	<u>hIgK</u>
Human IL-8	7 x 10 <sup>5</sup>	1 x 10 <sup>6</sup>	4 x 10 <sup>6</sup>
Human EGFR	4 x 10 <sup>5</sup>	7.5 x 10 <sup>5</sup>	3 x 10 <sup>6</sup>
Human TNF $\alpha$	4 x 10 <sup>5</sup>	4 x 10 <sup>5</sup>	3 x 10 <sup>5</sup>

FIGURE 18

# Composition of Anti-IL-8 Antibodies

Antibody	VH	D	JH	Vκ	Jκ
D1.1	4 - 34	21 - 10rc	JH3	018	Jκ3
K2.2	4 - 21	ir3rc	JH4	B3	Jκ3
K4.2	3 - 30	K1	JH4	018	Jκ4
K4.3	5 - 51	M5-a/M5-b	JH4	012	Jκ4

FIGURE 19

# IL-8 Antibodies: Heavy Chain Sequences

Human $\gamma 2$					ASTKQPSVTFPLAPCSRSTST
VH (D <sub>H</sub> /J <sub>H</sub> )					
4-34/4-21	CDR1	CDR2	CDR3		
	LSLTCAVYGGSFQYTNQSWIRQPEKQGLEWIGELNHSGSTNYNPBLKSRVTISVDTSKNQFSILKLSSTAAADTAVYICAR				
D1.1					
(D21-10rc/JH <sub>3</sub> )					ETPH AFDI WQQQTMTVVS
K4.2					
(DK1/JH <sub>4</sub> )					GCVAAPDY WQQQTMTVVS
5-51	CDR1	CDR2	CDR3		
	SLKISCKGSGYSTSYWIGWVRQMPQKQLEWGIYYPDSDTRYSPFQQTISADKKSISTAYLQWSSLKASDTAMYICAR				
K4.3					
(A <sub>1</sub> /A /JH )					
3-30					QEQDSFDY WQQQTMTVVS
K2.2					
(Dir3rc/JH <sub>4</sub> )					DRLGIFDY WQQQTMTVVS

FIGURE 20

# IL-8 Antibodies:

## Kappa Chain Sequences

Human CK					
Vk (J <sub>k</sub> )	CDR1	CDR2	CDR3		RTVAAPSVFI <del>PP</del> SD <del>EQ</del>
08	TITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVP <del>S</del> RFSGSGTDFTFTISSLQPEDIA <del>TY</del> YQQDNLP				
D1.1 (Jk3)	-----KF-S-F-----	-----GT-Y-----S-----L-----V-----D-----YTFOPGTKVDIK-----			
K4.2 (Jk4)	-----A-----V-----A-----A-----		-----H-----LTFGGGTXVEIK-----		
02	CDR1 TCRASQSISSYLNWYQQKPGKAPKLIYAASLSLQSGVPSRFSGSGTDFTLTISSLQPEDFAT <del>TY</del> YCCQSYSTP	CDR2 -----	CDR3 -----		
K4.3 (Jk4)	-----N-----F-----G-----E-----		-----N-----LTFGGGTXVEIK-----		
B3	CDR1 TINCKSSQSVLYSSNNKNYLNWYQQKPGQPPKLLITWASTRESGV <del>PD</del> RFSGSGTDFTLTISSLQAE <del>D</del> VAVYCCQYYSTP	CDR2 -----	CDR3 -----		
K2.2 (Jk3)	-----I-----S-----K-----		-----D-----FTFOPGTKVDIK-----		

FIGURE 21

# Anti-IL-8 Human Antibodies Block IL-8 Binding to Human Neutrophils

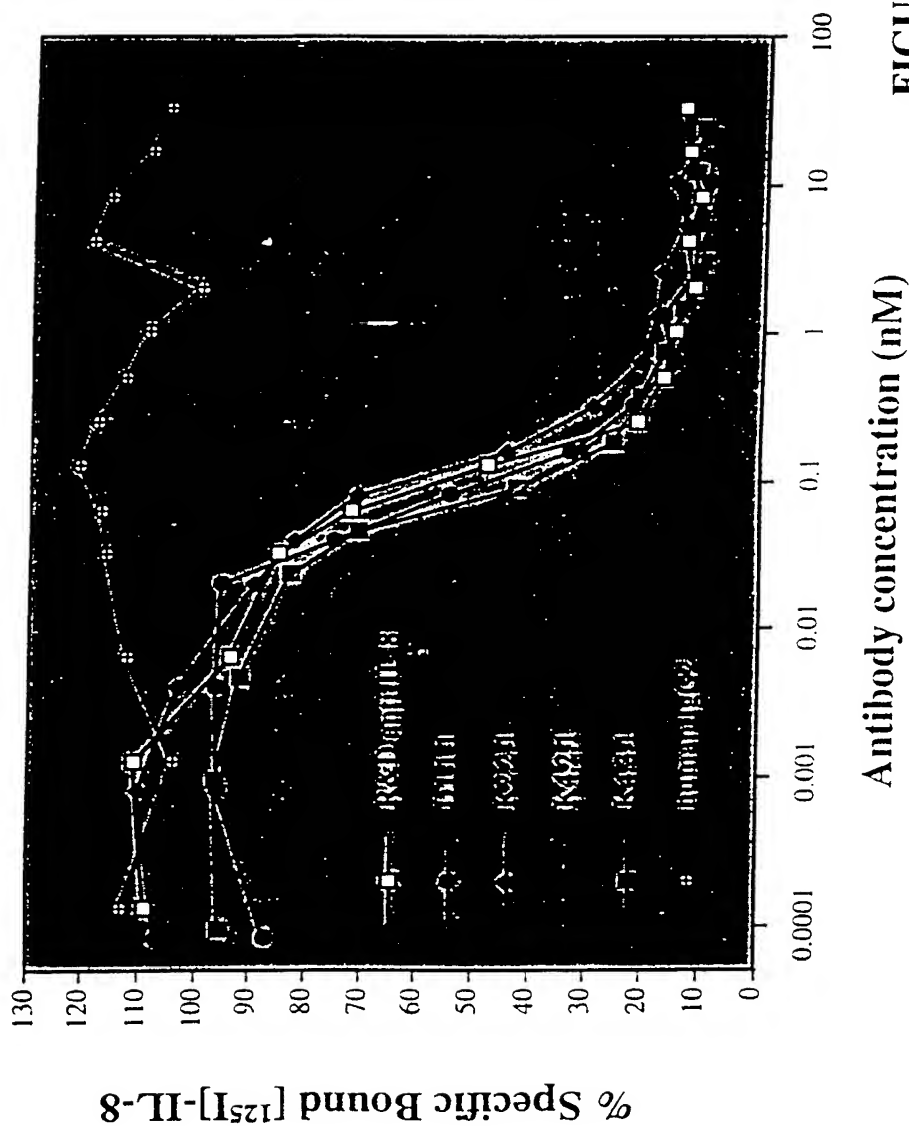


FIGURE 22

# Inhibition of CD11b Expression on Human Neutrophils by Anti-IL-8 Antibodies

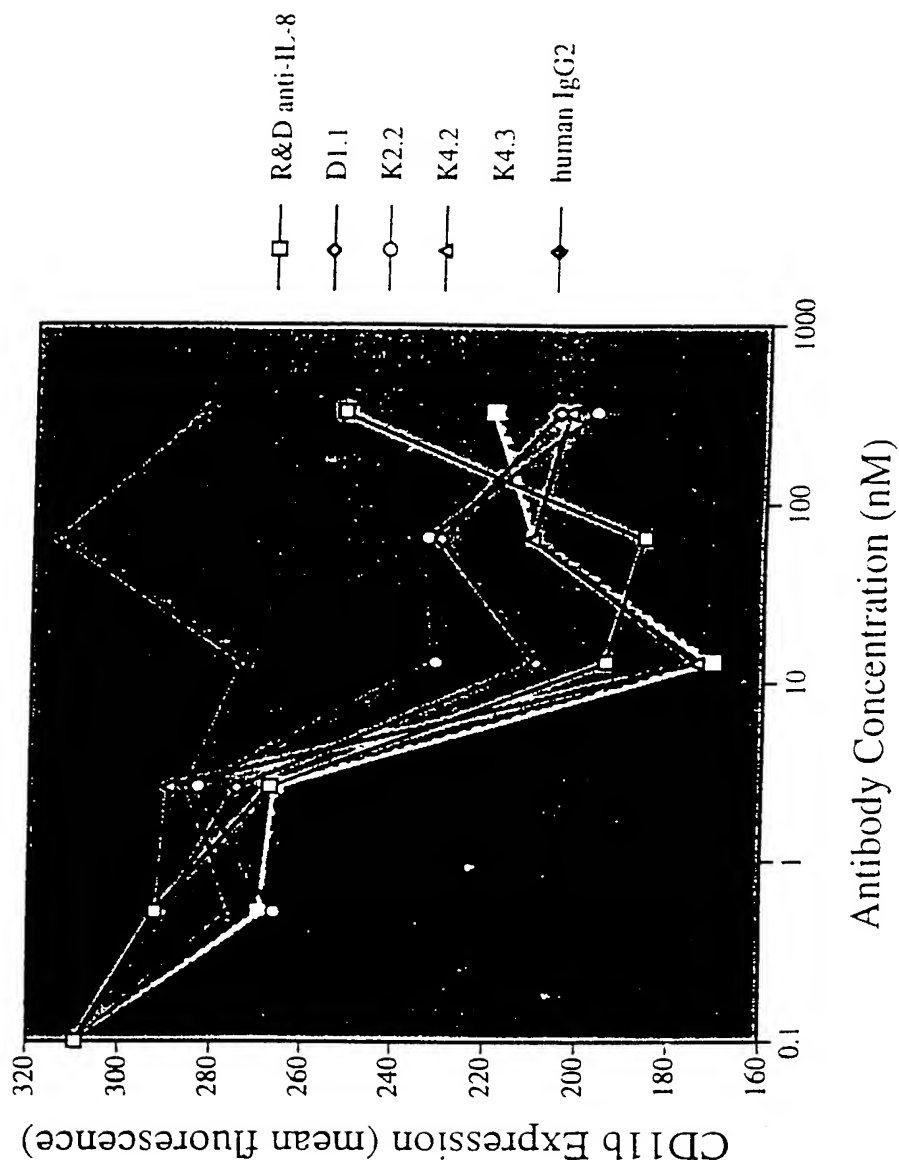


FIGURE 23



# Inhibition of IL-8 Induced $\text{Ca}^{+2}$ Influx by Anti-IL-8 Antibodies

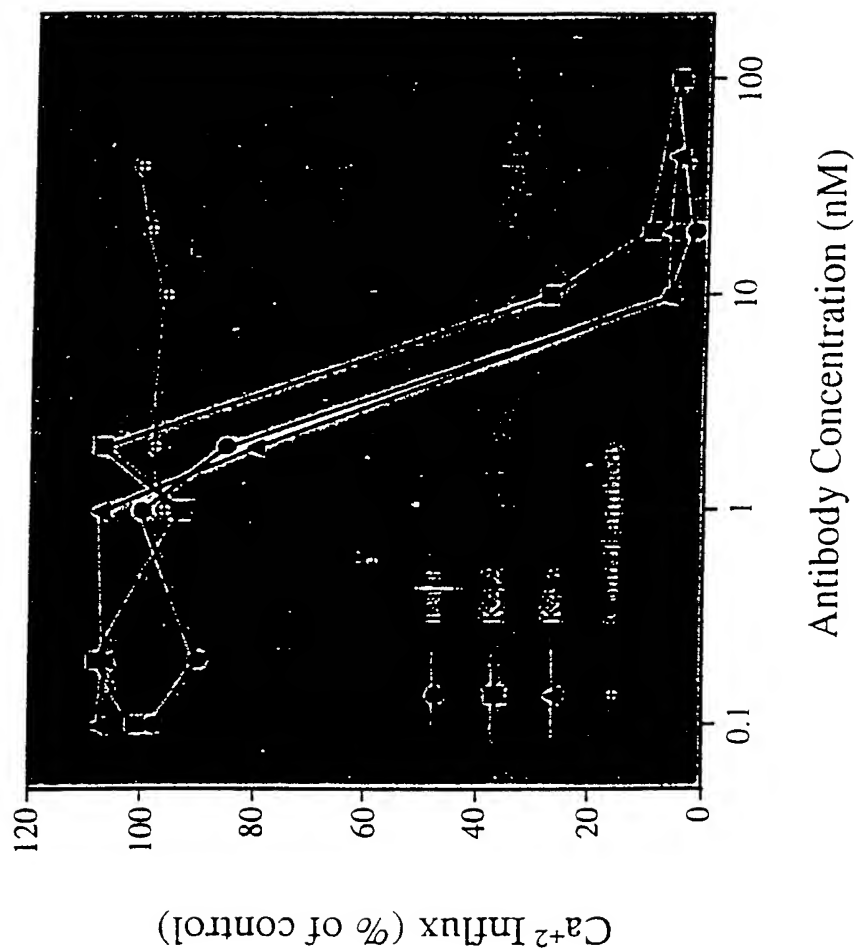


FIGURE 24

# Inhibition of IL8RB/293 Chemotaxis

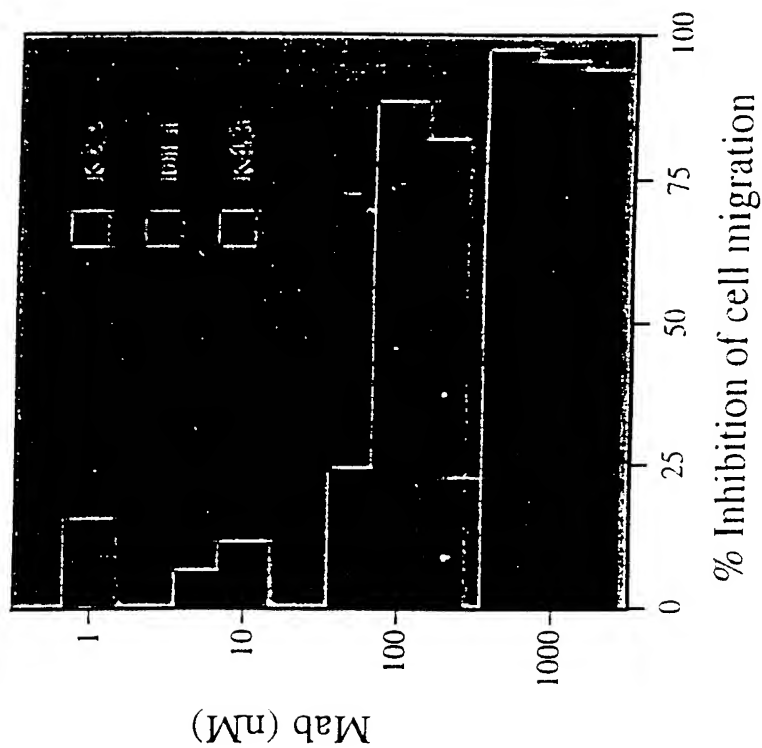


FIGURE 25

Wang and Oppenheim, NCI

# Rabbit Model of Human IL-8 Induced Skin Inflammation

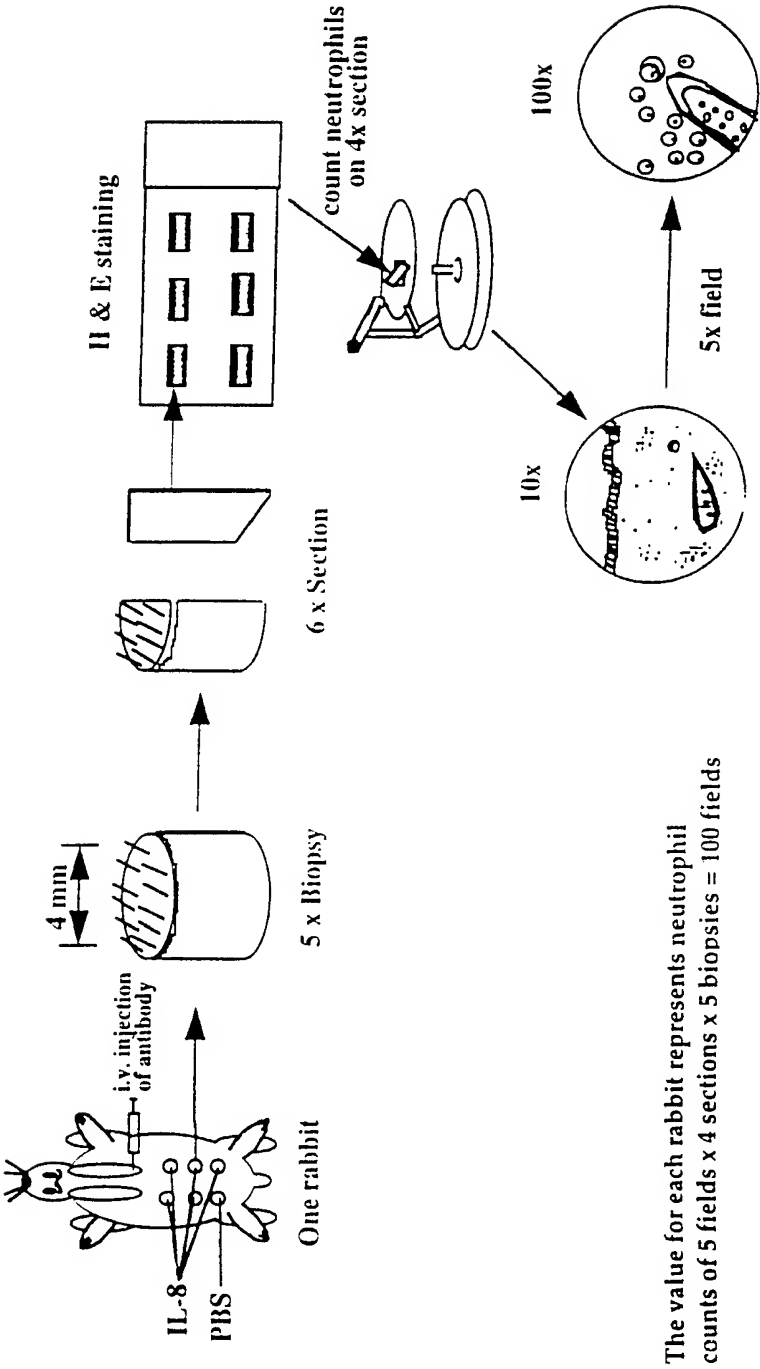


FIGURE 26

# Inhibition of IL-8-Induced Skin Inflammation in Rabbits by Anti-IL-8 Antibodies

Antibodies were administered i.v. single dose at 4mg/kg.

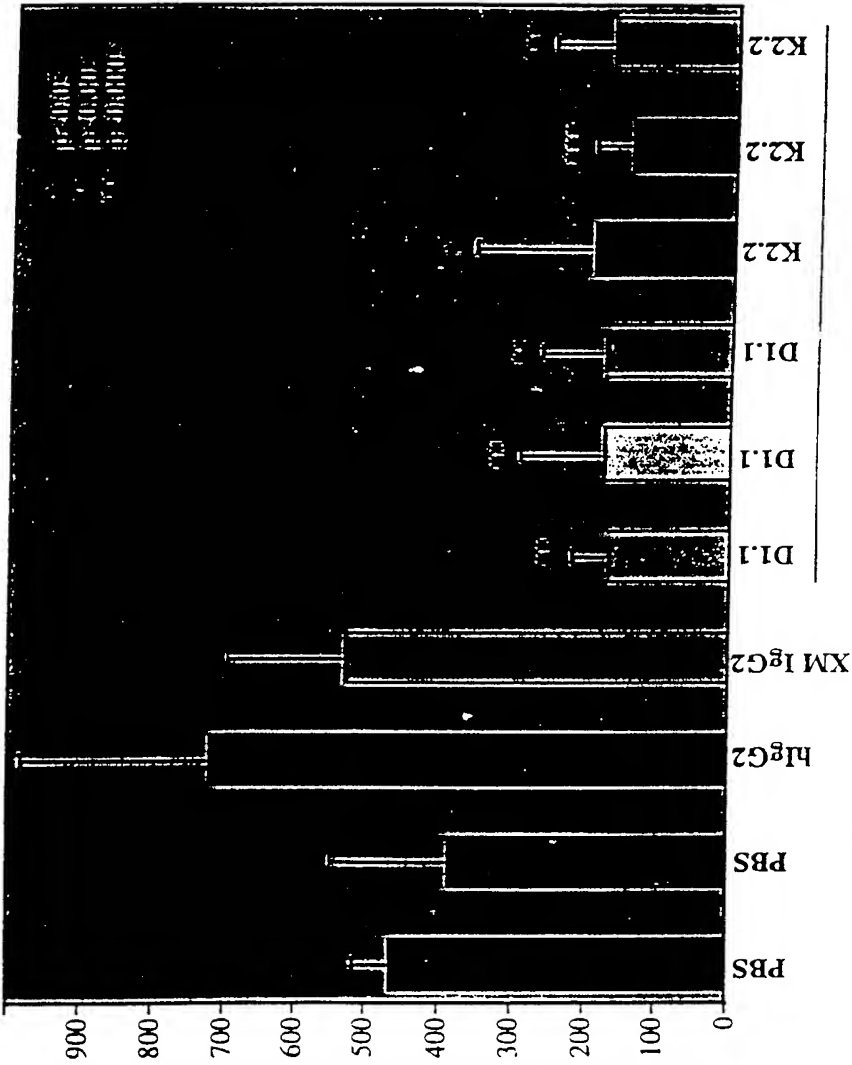


FIGURE 27

control human mAbs

# Inhibition of Angiogenesis Rat Corneal Pocket Model

WO 98/24893

28/35

PCT/US97/23091

<u>Implant</u>	Angiogenesis (# Corneas)	
	<u>Positive</u>	<u>Negative</u>
IL-8 alone	15	1
IL-8 + Control MAb	4	1
IL-8 + D1.1	0	6
IL-8 + K2.2	1	5
IL-8 + K4.3	2	4

FIGURE 28

# EGFR Antibody Gene Composition

Antibody	VH	D	JH	VK	JK
E1.1	4 - 31	2	JH5	018	JK4
E2.4	4 - 31	A1/A4	JH3	018	JK4
E2.5	4 - 31	XP1/21-10	JH4	018	JK4
E2.11	4 - 61	XP1/21-10	JH4	018	JK2

FIGURE 29

# EGFR Antibodies: Heavy Chain Sequences

Human $\gamma 2$					
$V_H (D_H/J_H)$	CDR1	CDR2	CDR3		ASTKQPSVTFPLAPCSRSTST
4-31	VSGGSISSGGYYWIRQPFCKGLEWIGYIYYSGSTYYNP	SLKSRVTISVDTSKNQFSLKLS	SSVTAADTA	VYYCAR	
E1.1 (D2/JH <sub>5</sub> )	-----N-D-----DC-----F--T-----		STVVPNGMFDP	WGQGYLVTVSS	
E2.4 (DA1/A4/JH <sub>3</sub> )	-----N-D-----S-----N-F-----L-----C-----		NIYTTQAFDI	WGQGTMTVTVSS	
E2.5 (DXF1/D21-10/JH <sub>4</sub> )	-----D---T-----N-----SM-I---E-----		KPVTGG EDYW	GQGTLLVTVSS	
4-61	CDR1	CDR2	CDR3		
	VSGGSISSGGYYWIRQPFCKGLEWIGYIYYSGSTYYNP	SLKSRVTISVDTSKNQFSLKLS	SSVTAADTA	VYYCAR	
E2.11 (DXF1/D21-10/JH <sub>4</sub> )	-----D-----HL-----N-----L-----		DYLTGSP FDY	WGQGTLLVTVSS	

FIGURE 30

# Anti-EGFR Human Antibodies Block EGF Binding to A431 Cells

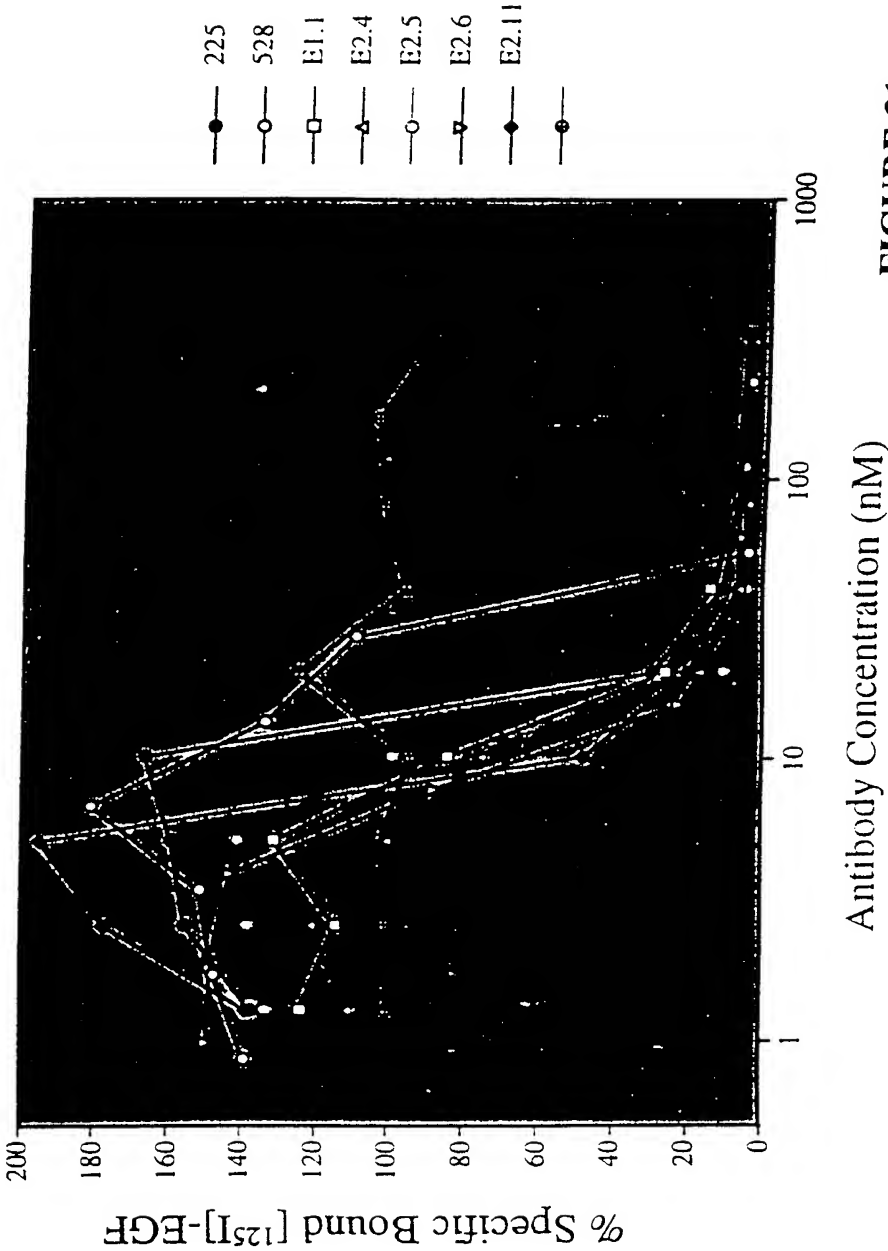


FIGURE 31



# Inhibition of EGF Binding to SW948 Cells by Anti-EGFR Human Antibodies

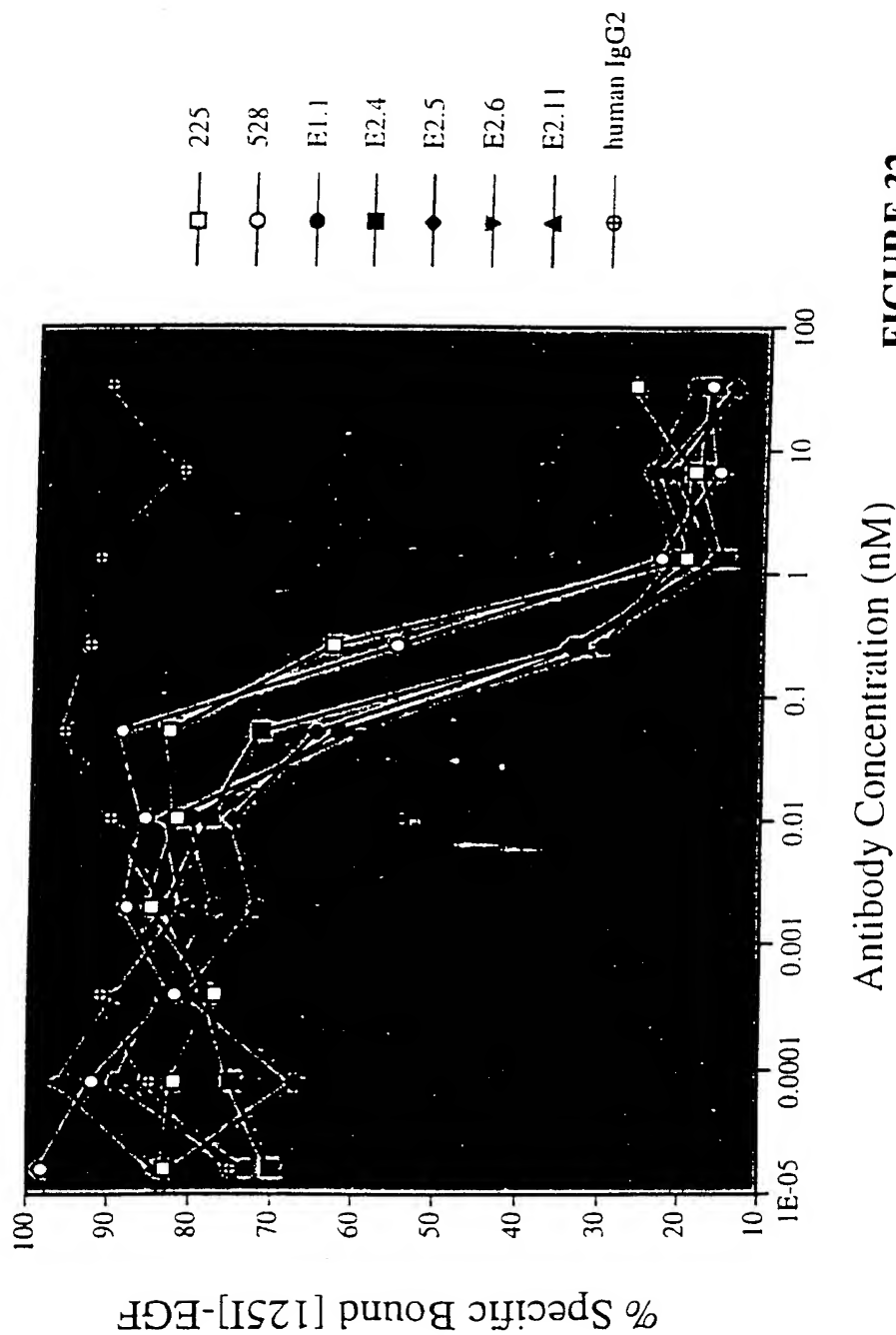


FIGURE 32

# Anti-EGFR Human Antibodies Inhibit SW948 Cell Growth

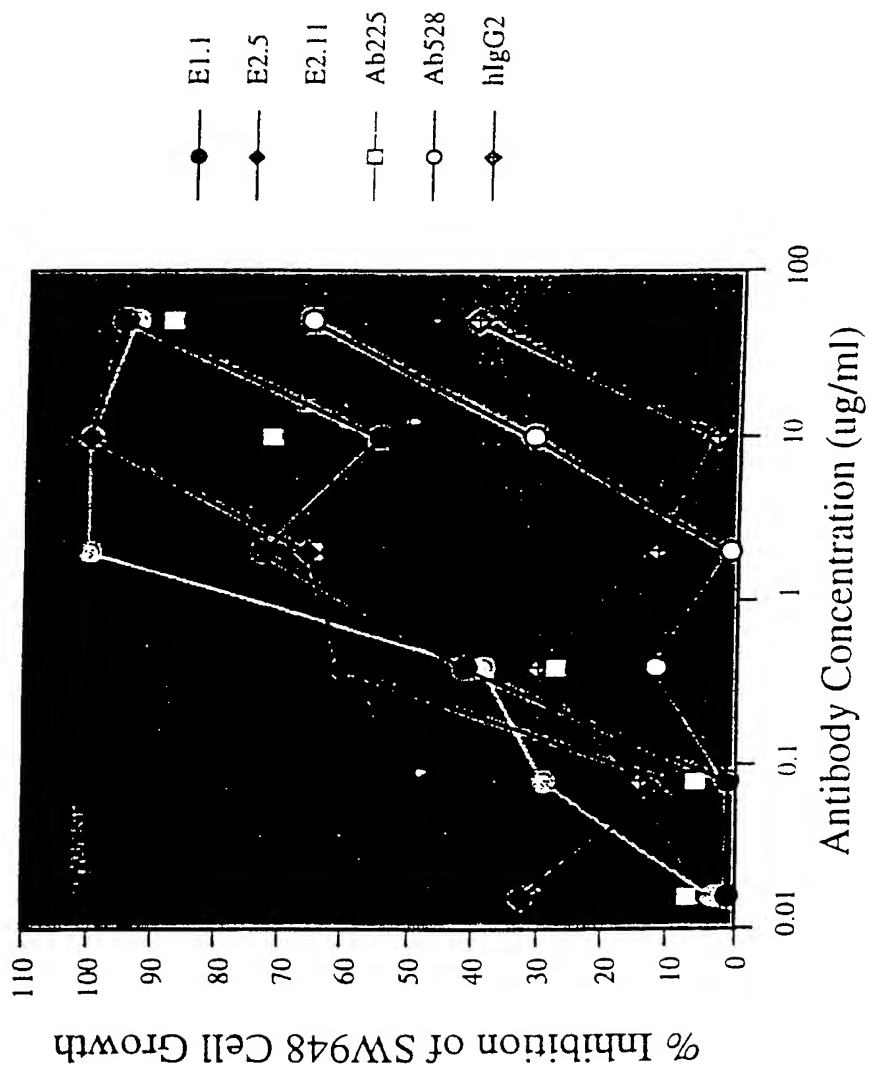
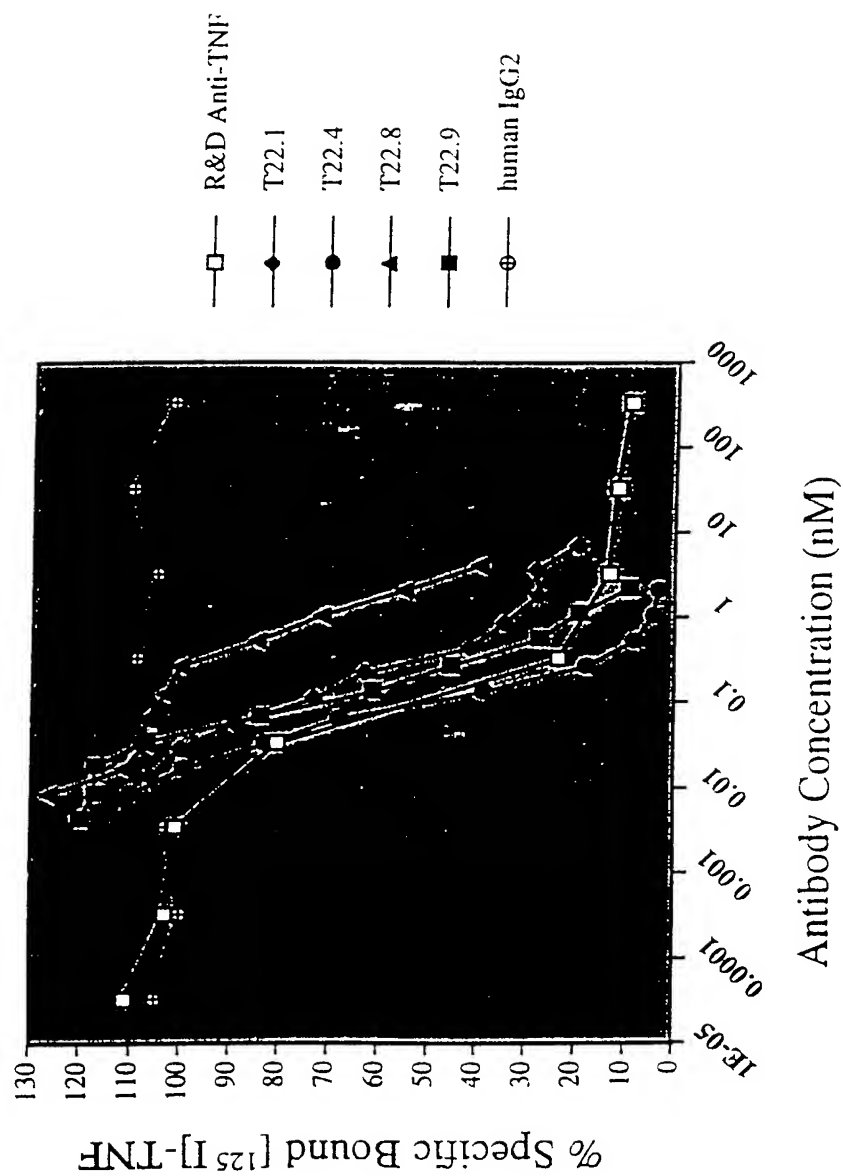


FIGURE 33

# Inhibition of TNF $\alpha$ Binding to U937 Cells by Anti-TNF $\alpha$ Human Monoclonal Antibodies

**FIGURE 34**

# anti-EGFR ANTIBODIES

## Kappa Light Chain V<sub>k</sub>(J<sub>k</sub>) Sequences

**E1.1**

TITCQASQDINNVLNWFQKPGKAPKVLIHDA  
SNLETGGPSRFSGSGSGTDFTTISGLQPE  
DIATYYCQQESLPLTFGGGKVEIKRTV  
AAPSVFI FPPSDEQ;

**E2.4**

TITCQASQDITIYLNWYQQKPGKAPKLL  
INDASSLETGVPLRFSGSGSGTDFTTIS  
SLQPEDIATYYCQQDHLPLTFGGGKVA  
IKRTV AAPSVFI FPPSDEQ;

**E2.5**

TITCQASQDISNYLNWYQQKPGKAPKLL  
IYDASNLETGVPSRFSGSGSGTDFTTIS  
SLQPEDVGTYVQQESLPCGFGGKLEIK  
RTV AAPSVFI FPPSDEQ;

**E2.11**

TITCQASQDISNYLNWYQQKPGKAPKLL  
INDASDLETGVPSRISGSGSGTDFTTIS  
NLQPEDIATYYCQQDSLPLTFGGGKVE  
IRRTV AAPSVFI FPPSDEQ;

**FIGURE 35**